



CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



**Junta
de Andalucía**

TESIS DOCTORAL

**Identificación por metagenómica
funcional de un operón completo de
biodesulfuración de dibenzotiofeno y
caracterización de su regulación**

Guadalupe Martín Cabello

TESIS DOCTORAL

Identificación por metagenómica funcional de un operón completo de biodesulfuración de dibenzotiofeno y caracterización de su regulación.

Memoria presentada por Guadalupe Martín Cabello para optar al grado de Doctora.

Sevilla, julio 2021.

El Director de la Tesis

La Doctoranda

Dr. Eduardo Santero Santurino

Guadalupe Martín Cabello

UNIVERSIDAD PABLO DE OLAVIDE
Dpto. Biología Molecular e Ingeniería Bioquímica

a pesar de todo y siempre,
Gracias a la vida que me ha dado tanto

Lo bueno de escribir estas líneas con mi edad, es que puedo hacer un poco de balance, porque mi trayectoria laboral no ha sido típica.

Desde pequeña me apasionaba la Biología, me parecía que lo más emocionante que se puede estudiar es la ciencia de la Vida, desde lo más pequeño, hasta lo más grande. Pero las circunstancias te van llevando por caminos distintos, te conviertes en un barco a la deriva, que sólo puede dejarse llevar. La pérdida de mi hermana me alejó de las ganas de saber, perdí el interés en conocer los procesos moleculares porque tratar de entender el porqué, resultaba demasiado doloroso.

Así que casi por casualidad, después de que Laura Tomás le diera mi curriculum a Belén (gracias por acordarte de mí) recibí una llamada de Edu un día de febrero del año 2005; me dijo “bueno, hasta ahora no has trabajado en biología molecular... pero necesitamos a alguien y vamos a probar contigo, a ver qué tal sale...” No sé qué tal ha salido, pero hemos llegado hasta aquí.

En fin, que después de tanto tiempo y tantas dificultades, después de estos 2 últimos años particularmente duros, parece que esto se acaba de verdad, y que ha llegado el momento de escribir los agradecimientos, espero no dejarme a nadie, porque hay tanta gente que me ha apoyado, que siento que a todos ellos les pertenece un trocito de esta Tesis.

Así que empezaré por Edu, muchas gracias por tu empeño en que yo hiciera el Doctorado, gracias por darme esta oportunidad de superarme a mí misma que me ha permitido recuperar las ganas de aprender, pero, además muchas gracias por toda tu ayuda durante la escritura, por el tiempo récord en que me has enviado todas las correcciones y por responderme a cada duda en cualquier momento.

En cuanto a los compañeros del laboratorio, todas las personas que han pasado por allí son importantes para mí, todas me han ayudado de alguna forma y de todas he aprendido algo; desde los primeros años con Vicente, Ana Hervás, Aroa..., la época en la que el laboratorio estaba siempre lleno con las tres Lauras, Ana Platero, Bea, Elena, Carlos... los últimos años con Inma, Rubén, Rocío, Eli, Yolanda, Magaly... seguro que me dejo a alguien.

Gracias a Eva, llegamos a micro casi a la vez, y en esos primeros meses en los que yo no sabía cómo se hacía nada, siempre estaba dispuesta a explicarme lo que tenía que hacer. A Belén, que fue la primera persona en decirme que debería plantearme hacer el Doctorado, a Paqui por tanta complicidad como siento que tengo con ella, y por sacarme una sonrisa cada vez que me veía con montones de tubos y matraces en la campana, a Fernan por estar siempre dispuesto a resolverme todo tipo de dudas, a Amando, por el ánimo que siempre me ha dado, a Inés que me enseñó a hacer la primera genoteca, a Nuria por estar siempre dispuesta a

escucharme cuando tengo un problema, a Aroa, no sólo por la portada, sino porque con ella aprendí a hacer betagales, a Marica, por trabajar con tanto entusiasmo, y por tantas curvas.

Mención especial merece algunas personas; Laura Terrón que inició este trabajo, pero sobre todo, por toda la ayuda con el análisis de secuencias, gracias por contestar a cada correo, por explicarme las mismas cosas una y otra vez, para mi eres mi codirectora. Rubén, por los buenos ratos que hemos pasado y por todas las cosas que he aprendido discutiendo contigo algún protocolo, ya sabes que me siento muy pequeña a tu lado, aunque juntos aprendimos que “cualquiera puede ser científico”. Te voy a echar mucho de menos. Por encima de todos, gracias a mi Elena, la mejor amiga, es imposible cuantificar todo lo que me has enseñado, desde purificar una proteína, hasta hacer un footprint, gracias por tantos ratos, buenos y malos, por tantos chismes compartidos, eres la persona más trabajadora que conozco, la única que siempre, está dispuesta a dejarlo todo y ayudar a los demás. Gracias.

Gracias a todos los que han estado en el laboratorio durante estos últimos meses, por la paciencia y la comprensión que me habéis mostrado tolerando mis ausencias.

De fuera del laboratorio tengo que empezar por mis padres. Gracias a su fe inquebrantable en mi, y a su apoyo incondicional, gracias a su entrega, esa forma de ser padres anteponiendo sus hijas a todo, incluso y, sobre todo, a sí mismos. Siempre me habéis pedido, “que fuera responsable y trabajadora” independientemente de los resultados que obtuviera, creo que he cumplido.

Gracias a mis hermanas, a las dos. A la mayor, que siempre estará conmigo, y a la pequeña, aunque hace mucho que no lo eres ¡te admiro tanto! por tus increíbles capacidades y por tu resiliencia. Pero, sobre todo, gracias por ser Hermana, con mayúscula y por ser la mejor tata, siempre.

A mis primos, por los buenos ratos compartidos y por la cobertura con mis hijos y mis padres.

Gracias a mi familia Guillén Cruz, por el cariño y el apoyo que me habéis dado a pesar del momento que estáis viviendo.

Y por supuesto a Guillermo y a mis hijos, gracias por ser, por estar, gracias por tanto... sin duda mi familia es el experimento del que estoy más orgullosa.

Gracias a los tres por apoyarme y creer en mi, aún si entender muy bien qué es lo que hago. Me enternece que los tres piensen en mi como “una científica”. Aunque yo me sienta como el último eslabón de la cadena, el eslabón más chiquitito e insignificante, cuando ellos piensan en mi, parece que soy un eslabón más importante, y tal vez, sólo por eso, todas las horas que la Tesis les ha robado de mi, es posible que merezcan la pena.

Gracias a todos.

| | |
|---|----|
| INTRODUCCIÓN | 1 |
| I. BIODESULFURACIÓN | 3 |
| II. REGULACIÓN | 19 |
| II.1. REGULACIÓN DE LOS GENES <i>dsz</i> | 19 |
| II.2. REGULACIÓN POR AZUFRE | 20 |
| III. METAGENÓMICA FUNCIONAL | 25 |
| OBJETIVOS | 31 |
| MATERIALES Y MÉTODOS | 35 |
| I. Estirpes, plásmidos y oligonucleótidos. | 37 |
| II. Técnicas microbiológicas: | 44 |
| II.1. Medios y condiciones de cultivo. | 44 |
| II.2. Conservación del material biológico. | 45 |
| II.3. Ensayos de crecimiento en distintas fuentes de azufre. | 45 |
| II.4. Ensayos de toxicidad de 2-hidroxibifenilo (2-HBF)..... | 46 |
| II.4.1. Ensayo de toxicidad de 2-HBF en cultivo líquido..... | 46 |
| II.4.2. Ensayo de toxicidad de 2-HBF en medio sólido. | 46 |
| II.5. Transferencia de ADN a estirpes bacterianas:..... | 46 |
| II.5.1. Transformación. | 46 |
| II.5.1.a. Transformación por choque térmico de <i>E. coli</i> | 46 |
| II.5.1.a.1. Preparación de células competentes. | 46 |
| II.5.1.a.2. Transformación. | 47 |
| II.5.1.b. Electrotransformación..... | 47 |
| II.5.1.b.1. Preparación de células electrocompetentes. | 47 |
| II.5.1.b.2. Electroporación. | 48 |
| II.5.2. Conjugación. | 48 |
| II.5.2.a. Conjugación triparental..... | 48 |
| II.6. Rastreo de actividad funcional. Búsqueda de clones con capacidad de crecimiento en DBT. | 49 |
| III. Técnicas bioquímicas:..... | 49 |
| III.1. Ensayos de actividad β -galactosidasa. | 49 |
| III.2. Manipulación de proteínas. | 50 |
| III.2.1. Electroforesis de proteínas en geles de poliacrilamida-SDS. | 50 |
| III.2.2. Superproducción de DszR. | 51 |
| III.2.3. Purificación de DszR..... | 52 |
| III.2.4. Cuantificación de proteínas. | 52 |
| III.3. Cuantificación de la producción de 2-hidroxibifenilo. | 52 |
| IV. Técnicas de biología molecular: | 53 |
| IV.1. Manipulación de ácidos nucleicos. | 53 |
| IV.1.1. Manipulación de ADN. | 53 |
| IV.1.1.a. Aislamiento de ADN plasmídico de <i>Escherichia coli</i> | 53 |
| IV.1.1.b. Electroforesis de ADN. | 54 |
| IV.1.1.c. Reacción en cadena de la polimerasa (PCR). | 54 |

| | | |
|------------|---|----|
| IV.1.1.d. | Aislamiento y purificación de fragmentos de ADN. | 55 |
| IV.1.1.e. | Cuantificación de preparaciones de ADN..... | 55 |
| IV.1.1.f. | Clonación de fragmentos de ADN. | 55 |
| IV.1.1.g. | Secuenciación y análisis de secuencias. | 56 |
| IV.1.2. | Manipulación de ARN. | 56 |
| IV.1.2.a. | Extracción de ARN total de <i>Escherichia coli</i> | 56 |
| IV.1.2.b. | Electroforesis de ARN. | 58 |
| IV.1.2.c. | Cuantificación de ARN. | 58 |
| IV.1.2.d. | Retrotranscripción de ARN..... | 58 |
| IV.2. | Determinación del extremo 5' de ARN mensajeros por el método de extensión desde el cebador (<i>primer extension</i>)..... | 58 |
| IV.2.1. | Marcaje de los oligonucleótidos con γ [32P]ATP..... | 58 |
| IV.2.2. | Reacción de extensión de la retrotranscriptasa..... | 59 |
| IV.2.3. | Análisis en gel de secuenciación..... | 60 |
| IV.3. | PCR cuantitativa (qPCR)..... | 60 |
| IV.4. | Ensayo de transcripción in vitro..... | 60 |
| IV.5. | Análisis de interacciones ADN-proteína..... | 61 |
| IV.5.1. | Marcaje de sondas de ADN..... | 61 |
| IV.5.2. | Ensayo de retardo de movilidad en gel..... | 61 |
| IV.5.3. | Ensayo de protección frente a digestión por DNasa I (<i>footprinting</i>). | 62 |
| RESULTADOS | | 65 |
| I. | Caracterización de los genes catabólicos <i>dszEABC</i> | 67 |
| I.1. | Aislamiento y caracterización del crecimiento del clon UPO112..... | 67 |
| I.1.1. | Búsqueda de un clon capaz de conferir a <i>E. coli</i> la capacidad de crecer en DBT como fuente de azufre. | 67 |
| I.1.2. | Caracterización del crecimiento del clon UPO112..... | 68 |
| I.2. | Análisis de secuencias de UPO112. | 71 |
| I.2.1. | Asignación de <i>orfs</i> | 71 |
| I.2.2. | Comparativa con UPO21 y <i>Comamonadaceae bacterium</i> | 76 |
| I.2.3. | Isla genómica..... | 84 |
| II. | Caracterización de la regulación del operón..... | 87 |
| II.1. | Caracterización de la expresión del operón <i>dszEABC</i> de UPO21..... | 87 |
| II.1.1. | Expresión propia y desde P_{sal} . Regulación postranscripcional. | 87 |
| II.1.2. | Determinación del inicio de transcripción desde P_{dsz} | 89 |
| II.1.3. | Mutantes de delección ΔP_{dsz} . Expresión desde P_{sal} | 91 |
| II.1.4. | Estructura secundaria del ARNm..... | 92 |
| II.2. | Caracterización del papel de DszR. | 93 |
| II.2.1. | Actividad de DszR <i>in vivo</i> | 93 |
| II.2.2. | Caracterización <i>in vitro</i> de la actividad de DszR..... | 95 |
| II.2.2.a. | Purificación de DszR. | 95 |
| II.2.2.b. | Ensayos de retardo de movilidad en gel. | 96 |
| II.2.2.c. | Identificación de los sitios de unión de DszR. | 97 |
| II.2.2.d. | Ensayos de Transcripción <i>In Vitro</i> | 99 |

| | | |
|----------|--|-----|
| II.3. | Regulación por fuente de azufre. | 100 |
| II.3.1. | Expresión de <i>dszE</i> _{UPO112} :: <i>lacZ</i> | 100 |
| II.3.2. | Transcripción de <i>dszE</i> y <i>dszR</i> dependiente de fuente de azufre. | 101 |
| II.4. | Papel de <i>dszHR</i> en la regulación por fuente de azufre. | 102 |
| II.4.1. | Subclonación de <i>dszH-dszR</i> _{wt} | 102 |
| II.4.2. | Regulación de <i>dszE</i> :: <i>lacZ</i> por <i>dszHR</i> _{wt} en limitación de azufre. | 103 |
| II.4.3. | Expresión de <i>dszEABC</i> en limitación de nitrógeno. | 104 |
| II.4.4. | Deleción interna en <i>dszH</i> | 105 |
| II.4.5. | Expresión de <i>dszR</i> desde promotores heterólogos. | 107 |
| III. | Construcción de casetes de biodesulfuración de DBT de amplio espectro. | 111 |
| III.1. | Toxicidad de 2-HBF en <i>E. coli</i> y en <i>Pseudomonas</i> | 111 |
| III.1.1. | Toxicidad en medio líquido. | 111 |
| III.1.2. | Toxicidad en medio sólido. | 112 |
| III.2. | Diseño y construcción de casetes catabólicos de amplio espectro. | 113 |
| III.3. | Validación en <i>E. coli</i> y <i>Pseudomonas</i> | 114 |
| III.3.1. | Expresión desde su propio promotor. | 114 |
| III.3.2. | Expresión desde el sistema <i>alkS-P_{alkB}</i> | 116 |
| | DISCUSIÓN | 119 |
| I. | Características diferenciadoras de los genes <i>dsz</i> en UPO21 y UPO112. Importancia de los estudios de metagenómica funcional. | 121 |
| II. | Transferencia horizontal de los operones <i>dszEABC</i> de UPO21 y UPO112. | 124 |
| III. | Regulación de la expresión de <i>dszEABC</i> | 125 |
| IV. | Función de <i>DszH</i> | 129 |
| V. | Casetes catabólicos de biodesulfuración de DBT. | 131 |
| | CONCLUSIONES | 135 |
| | ANEXO | 139 |
| | BIBLIOGRAFÍA | 143 |

ÍNDICE DE TABLAS

| | |
|---|-----|
| Tabla 1: Estirpes empleadas en este trabajo | 33 |
| Tabla 2: Plásmidos empleados en este trabajo | 33 |
| Tabla 3: Oligonucleótidos empleados en este trabajo | 39 |
| Tabla 4: DO ₆₀₀ tras 48 h de crecimiento en DBT como única fuente de azufre de los candidatos obtenidos del rastreo funcional..... | 62 |
| Tabla 5: Resultados de comparar las proteínas codificadas por cada <i>orf</i> de UPO112 con las bases de datos: “non-redundant protein sequences” (nr) (blastp), y UniProtKB/Swiss-Prot (swissprot)..... | 66 |
| Tabla 6: Comparativa del tamaño de los genes y proteínas de UPO21 vs UPO112..... | 70 |
| Tabla 7: Comparativa regiones intergénicas UPO21 vs UPO112 vs <i>Comamonadaceae</i> <i>bacterium</i> | 70 |
| Tabla 8: Resultados de comparar <i>orfs</i> de UPO21 vs <i>orfs</i> UPO112 (blastn/blastp)..... | 71 |
| Tabla 9: Resultados de comparar <i>orfs</i> de UPO112 vs <i>Comamonadaceae bacterium</i> (blastn)..... | 71 |
| Tabla 10: Organización génica de las <i>orfs</i> de UPO112 y de <i>orfs</i> de <i>Candidatus</i> <i>Macondimonas diazotrophica</i> , con las que presentan homología..... | 79 |
| Tabla 11: Producción de 2-HBF con pMPO1566 o pSEVA221..... | 110 |
| Tabla 12: Producción de 2-HBF con pMPO1570 o pSEVA429..... | 111 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| Figura 1: Representación esquemática de rutas de desulfuración de compuestos organosulfurados..... | 4 |
| Figura 2: Representación esquemática de la regulación transcripcional de los genes SSI, en <i>E. coli</i> , <i>P. putida</i> DS1 y <i>P. aeruginosa</i> | 20 |
| Figura 3: Curvas de crecimiento del clon UPO112 en medio mínimo con DBT como única fuente de azufre y adición de diferentes inductores..... | 63 |
| Figura 4: Curvas de crecimiento del clon UPO112 en medio mínimo con diferentes fuentes de azufre..... | 64 |
| Figura 5: Curvas de crecimiento del clon UPO112 en medio mínimo con DBT y DMDBT como fuentes de azufre..... | 64 |
| Figura 6: Producción de índigo por parte de los clones UPO21 y UPO112, en medio LB..... | 65 |
| Figura 7: Esquema de la organización génica de los fósidos UPO21 y UPO112..... | 70 |
| Figura 8: Alineamiento múltiple de DszE _{UPO21} , DszE _{UPO112} y NADPH-dependent FMN reductase de <i>Comamonadaceae bacterium</i> | 72 |
| Figura 9: Árbol filogenético de DszE _{UPO112} , DszE _{UPO21} NADPH-dependent FMN reductase de <i>Comamonadaceae bacterium</i> junto con sulfonato monooxigenasas alifáticas y reductasas de DBT..... | 73 |
| Figura 10: Alineamiento múltiple de SfnR de <i>P. Fluorescens</i> , DszR _{UPO112} y DszR _{UPO21} | 75 |
| Figura 11: Árbol filogenético de DszR _{UPO112} y DszR _{UPO21} junto con otros activadores transcripcionales dependientes de σ^{54} | 76 |
| Figura 12: Alineamiento múltiple de la proteína codificada por <i>gen11</i> _{UPO21} vs DszH _{UPO112} vs la alfa beta hidrolasa de <i>Propionovibrio dicarboxylicus</i> | 78 |
| Figura 13: Esquema y actividad betagalactosidasa de las fusiones transcripcionales y traduccionales a <i>lacZ</i> de <i>dszE</i> _{UPO21} , <i>dszEA</i> _{UPO21} y <i>dszEABC</i> _{UPO21} | 82 |
| Figura 14: Alineamiento de la región promotora de <i>dszE</i> _{UPO21} y <i>dszE</i> _{UPO112} | 83 |
| Figura 15: Determinación del inicio de transcripción de <i>dszE</i> | 84 |
| Figura 16: Esquema y actividad betagalactosidasa de las fusiones traduccionales a <i>lacZ</i> con <i>Pdsz</i> deletado..... | 85 |
| Figura 17: Estructuras secundarias predichas de la región líder de <i>dszE</i> de UPO21..... | 87 |
| Figura 18: Actividad de DszR _{UPO21} <i>in vivo</i> | 88 |
| Figura 19: Purificación de DszR _{UPO21} | 89 |
| Figura 20: Retardo en gel con DszR _{UPO21} y la región aguas arriba <i>dszE</i> de UPO21..... | 90 |
| Figura 21: Región promotora de <i>dszE</i> . Ensayo de protección frente a DNasaI..... | 92 |
| Figura 22: Transcripción <i>in vitro</i> desde <i>Pdsz</i> | 93 |
| Figura 23: Actividad betagalactosidasa de la fusión transcripcional <i>dszE</i> _{UPO112} :: <i>lacZ</i> en UPO112, con diferentes fuentes de azufre..... | 95 |
| Figura 24: Niveles de transcripción de <i>dszE</i> y <i>dszR</i> de UPO112 determinada por RT-qPCR, en diferentes fuentes de azufre..... | 96 |
| Figura 25: Actividad betagalactosidasa de la fusión transcripcional <i>dszE</i> _{UPO112} :: <i>lacZ</i> , con el plásmido con los genes <i>dszHR</i> en fuentes de azufre..... | 97 |
| Figura 26: Actividad betagalactosidasa de la fusión transcripcional <i>dszE</i> _{UPO112} :: <i>lacZ</i> , con el plásmido con los genes <i>dszHR</i> diferentes fuentes de nitrógeno y de azufre..... | 98 |
| Figura 27: Actividad betagalactosidasa de la fusión transcripcional <i>dszE</i> _{UPO112} :: <i>lacZ</i> con el plásmido $\Delta dszH_1$, en diferentes fuentes de azufre..... | 99 |

| | |
|---|-----|
| Figura 28: Niveles de transcripción de <i>dszR</i> , determinados por RT-qPCR con cDNA de <i>dszHR_{wt}</i> y $\Delta dszH_1$ en taurina como fuente de azufre..... | 100 |
| Figura 29: Complementación de $\Delta dszH_1$ con el plásmido que lleva $\Delta P_{dsz}dszH$ bajo P_{tac} | 101 |
| Figura 30: Actividad betagalactosidasa de la fusión transcripcional <i>dszE_{UPO112}::lacZ</i> con ΔP_{dszHR} y $\Delta P_{dszHR} \Delta dszH_2$, ambos bajo P_{sal} , en diferentes fuentes de azufre..... | 102 |
| Figura 31: Actividad betagalactosidasa de la fusión transcripcional <i>dszE_{UPO112}::lacZ</i> con ΔP_{dszHR} y $\Delta P_{dszHR} \Delta dszH_2$, ambos bajo P_{tac} , en diferentes fuentes de azufre..... | 103 |
| Figura 32: Ensayo de toxicidad en medio líquido. MPO554 NaI ^R / UPO112 con diferentes fuentes de azufre, y diferentes concentraciones de 2-HBF..... | 106 |
| Figura 33: Ensayo de toxicidad en medio sólido; estirpes MPO554 NaI ^R , KT2440 y DOT T1E, con sulfato como fuente de azufre, y diferentes concentraciones de 2-HBF..... | 107 |
| Figura 34: Mapa genético de los plásmidos pMPO1566 y pMPO1570..... | 108 |
| Figura 35: Curvas de crecimiento de MPO554 NaI ^R , KT2440 y DOT T1E con pMPO1566, en medio mínimo con DBT como fuente de azufre..... | 109 |
| Figura 36: Curvas crecimiento de KT2440, DOT T1E y MPO406 con pMPO1570, en medio mínimo con DBT como fuente de azufre..... | 111 |
| Figura 37: Representación esquemática del modelo propuesto de regulación transcripcional del operón <i>dszEABC</i> | 121 |

ABREVIATURAS

ADN: ácido desoxirribonucleico
Abs_x: densidad óptica medida a x nm
Ap: ampicilina; Ap^R: resistente a ampicilina
Ara: arabinosa
ARN: ácido ribonucleico
ARNm: ARN mensajero
ARNp: ARN polimerasa
ATP: adenosín trifosfato
BAC: cromosoma artificial bacteriano (*bacterial artificial chromosome*)
BDS: biodesulfuración
BSA: Albúmina de suero bovino (*Bovine Serum Albumin*)
°C: grados Celsius
Cb: carbenicilina; Cb^R: resistente a carbenicilina
cDNA: ADN complementario
Cm: cloranfenicol; Cm^R: resistente a cloranfenicol
cm: centímetro
cpm: cuentas por minuto
DBT: dibenzotiofeno
DCPK: diciticlopropilcetona
DEPC: dietilpirocarbonato
DMDBT: 4,6-dimetil-dibenzotiofeno
DMS: sulfuro de dimetilo
DMSO: dimetil sulfóxido
DMSO₂: dimetil sulfona
dNTP: desoxinucleótido trifosfato
D.O._x: densidad óptica medida a x nm
DTT: ditiotreitolo
EDO: extradiol dioxigenasa
EDTA: ácido etilendiaminotetracético
et al.: y otros
FMN: flavín mononucleótido
Fw: *Forward*
g: aceleración de la gravedad terrestre
g: gramos
Gm: gentamicina; Gm^R: resistente a gentamicina
h: hora
2-HBF: 2-hidroxibifenilo
HBFS: 2-hidroxibifenilo-2-sulfonato
HDS: hidrodeshulfuración
IC₅₀: concentración media inhibitoria
IPTG: isopropil-β-D-tiogalactósido
IVT: transcripción in vitro (*in vitro transcription*)
kb: kilo pares de bases
kDa: kilodalton
Km: kanamicina; Km^R: resistente a kanamicina
K_m: constante de afinidad
Kv: kilovoltio
l: litro
LB: medio Luria-Bertani
M: molar

m: mili
M&M: materiales y métodos
M9: medio mínimo M9
M9/A9: medio mínimo M9 modificado, con solución de sales A9
min: minutos
MSA: metanosulfonato
n: nano
NAD: nicotinamida-adenín dinucleótido
Nal: ácido nalidíxico, Nal^R: resistente a ácido nalidíxico
NCBI: *National Center for Biotechnology Information*
NEB: New England Biolabs
NNDMF: NN-dimetil formamida
ONPG: ortho-Nitrophenyl-β-galactoside
orf: pauta abierta de lectura (*open reading frame*)
PAGE: electroforesis en gel de poliacrilamida (*Poly-Acrylamide Gel Electrophoresis*)
pb: pares de bases
PCR: reacción en cadena de la polimerasa (*polymerase chain reaction*)
RBS: sitio de unión de ribosomas (*ribosome binding site*)
Rif: Rifampicina; Rif^R: Rifampicina resistente
r.p.m.: revoluciones por minuto
RT-qPCR: PCR cuantitativa
Rv: *reverse*
s: segundo
Sal: salicilato
SD: Secuencia *Shine Dalgarno*
SDS: dodecil sulfato sódico
sitio *nut*: sitio de utilización de la proteína antiterminadora N (*N-utilization*)
SSI: inducción por limitación de azufre (*sulfur starvation induced*)
SSR: respuesta a limitación de azufre (*sulfur starvation response*)
Str: Streptomicina; Str^R: Streptomicina resistente
t: tiempo
TAE: Tris-Acetato-EDTA
Tau: Taurina (ácido 2-aminoetanosulfónico)
TBE: Tris-Borato-EDTA
Tc: tetraciclina; Tc^R: tetraciclina resistente
TE: Tris-EDTA
TEMED: N,N,N',N'-tetrametiletilendiamina
TER: Tris-EDTA-RNasa
TSS: sitio de inicio de la transcripción (*transcription start site*)
P_m: peso molecular
P_{T7}: promotor del gen 10 del fago T7
P_{tac}: promotor tac
P_{dsz}: promotor de los genes *dsz*
P_{dszHR}: promotor de los genes *H* y *R*
p/v: peso/volumen
U. M.: unidades Miller
U. A.: unidades arbitrarias
UV: ultravioleta
v/v: volumen/volumen
wt: estirpe silvestre (*wt*)
X-gal: (5-bromo-4-cloro-3-indolil-β-D-galactopiranosido)
λ: longitud de onda

Δ : deleción del gen

μ : micro

σ : factor sigma

INTRODUCCIÓN

I. BIODESULFURACIÓN

El azufre es un componente natural de los crudos de petróleo, en los que supone entre el 0,05 y el 10% de su composición. La quema de combustibles fósiles libera dióxido de azufre, que contribuye en gran medida a la lluvia ácida, la cual envenena bosques y lagos, produce corrosión en edificios y altera el equilibrio natural de los elementos químicos en el medio ambiente, afectando la diversidad biológica de los ecosistemas. Además, las emisiones totales de partículas, que son una grave fuente de contaminación atmosférica, son directamente proporcionales a la cantidad de azufre en el combustible diésel, y según la Agencia de Protección Ambiental de los Estados Unidos (USEPA), aproximadamente el 2% del azufre del combustible diésel se convierte en emisiones directas de partículas (Mohebbi & Ball, 2016).

La legislación ambiental es cada vez más estricta, y exige combustibles, con niveles de azufre ultra bajos, lo que supone cantidades de azufre inferiores a 15ppm (Martínez *et al.*, 2017). Por lo tanto, es necesario desulfurar los combustibles. En la industria petroquímica, este proceso se lleva a cabo fundamentalmente mediante hidrodesulfuración (Wang *et al.*, 2017), en la cual las distintas fracciones de petróleo se mezclan con hidrógeno de alta pureza y se pasan por un catalizador con determinadas condiciones de presión y temperatura, para romper los enlaces carbono-azufre y convertir el azufre orgánico en hidrocarburos y en H_2S ; a mayor número de enlaces aromáticos presentes, más complicado es el proceso. Debido al elevado coste del proceso, con altos requerimientos de presión y temperatura, así como la baja disponibilidad de hidrógeno es necesario desarrollar procesos alternativos (Kilbane II, 2016).

Los compuestos orgánicos de azufre que se encuentran en los crudos de petróleo se pueden clasificar en: no heterocíclicos (tioles, sulfuros y disulfuros) y heterocíclicos (compuestos orgánicos de azufre cíclicos o multicíclicos condensados). Los compuestos heterocíclicos que contienen azufre se encuentran entre los contaminantes ambientales más potentes (Mohebbi & Ball, 2016). El dibenzotiofeno, o DBT es una molécula heterocíclica aromática de azufre, cuyos derivados con grupos alquilo, que son los compuestos orgánicos de azufre más abundantes en las fracciones pesadas del petróleo, son recalcitrantes a la desulfuración química, (Wang *et al.*, 2017), por lo que, debido a las actuales normativas ambientales, el proceso de biodesulfuración (BDS), en el que se emplean microorganismos, libres o inmovilizados, sus enzimas o extractos celulares como biocatalizadores, para eliminar el azufre de los combustibles

fósiles, debe ser tenido en consideración, si no como alternativa total si al menos como complemento a la desulfuración química. Una opción interesante, es incorporar la biodesulfuración después del hidrotratamiento, para eliminar así buena parte de los compuestos orgánicos de azufre recalcitrantes a la hidrodessulfuración, haciendo el proceso de eliminación de azufre más barato y eficiente. (Kilbane II, 2006; Kilbane II, 2016). La BDS combina una alta eficiencia de eliminación de azufre, gran selectividad de sustrato y baja generación de subproductos indeseables. Este proceso, ha sido estudiado ampliamente en los últimos 25 años, utilizando como compuesto modelo el DBT. Las bacterias desulfuradoras de compuestos heterocíclicos de azufre están muy extendidas en diferentes entornos y ubicaciones geográficas, lo que sugiere que la desulfuración es una estrategia importante de obtención de azufre para dichas especies. (Martínez *et al.*, 2017).

El DBT puede ser metabolizado por muchos microorganismos. Las vías metabólicas para su utilización se pueden clasificar en cuatro categorías: (1) oxidación de azufre, catalizada por lacasas de hongos y dioxigenasas aromáticas de bacterias, que convierten DBT en Dibenzotiofeno-5,5-dióxido (DBTO₂) y / o derivados hidroxilados, en esta vía no se libera el azufre. (2) Rotura carbono-carbono, o “vía Kodama”, realizada por algunas Gram negativas como *Rhizobium meliloti*, en la cual el azufre no se elimina de manera selectiva, sino que se rompe el esqueleto carbonado, reduciendo el valor calórico del fuel, por lo que no es de interés para la industria. (3) Rotura C-C y C-S, realizada por algunos microorganismos, como *Brevibacterium sp.* Es una vía poco extendida y poco caracterizada, en la que se rompe el esqueleto carbonado, por lo que no es de interés. (4) Rotura específica de los enlaces C-S. La más estudiada es la vía aerobia o “vía 4S”, en la cual se mantiene el esqueleto carbonado y la molécula no pierde valor como combustible, lo que hace que sea útil en la biodesulfuración del petróleo. Se trata de una vía enzimática en 4 pasos, que produce la liberación de azufre en forma de bisulfito, el cual puede ser empleado por la bacteria como fuente de azufre. Esta vía se encuentra ampliamente distribuida tanto en bacterias Gram positivas, donde se ha estudiado en más profundidad, como en Gram negativas. (Kilbane, J. J. 1989) (Xu, *et al.*, 2006) (Abin-Fuentes *et al.*, 2013) Además, existen ciertos microorganismos capaces de realizar una vía “4S extendida”, cuyo producto final en lugar de 2-HBF, es 2-metoxibifenilo o bifenilo, pero los genes responsables de estas rutas aún no han sido caracterizados, (Martínez *et al.*, 2017). Por otra parte, además de la vía 4S, existen ciertas bacterias Gram negativas reductoras de sulfato, como algunos *Desulfovibrio* capaces de eliminar específicamente el azufre del benzotiofeno y del DBT en condiciones anaeróbicas. Pero este proceso es muy poco conocido, lo que hace que, al menos por el momento, sea comercialmente inviable. (Martínez *et al.*, 2017).

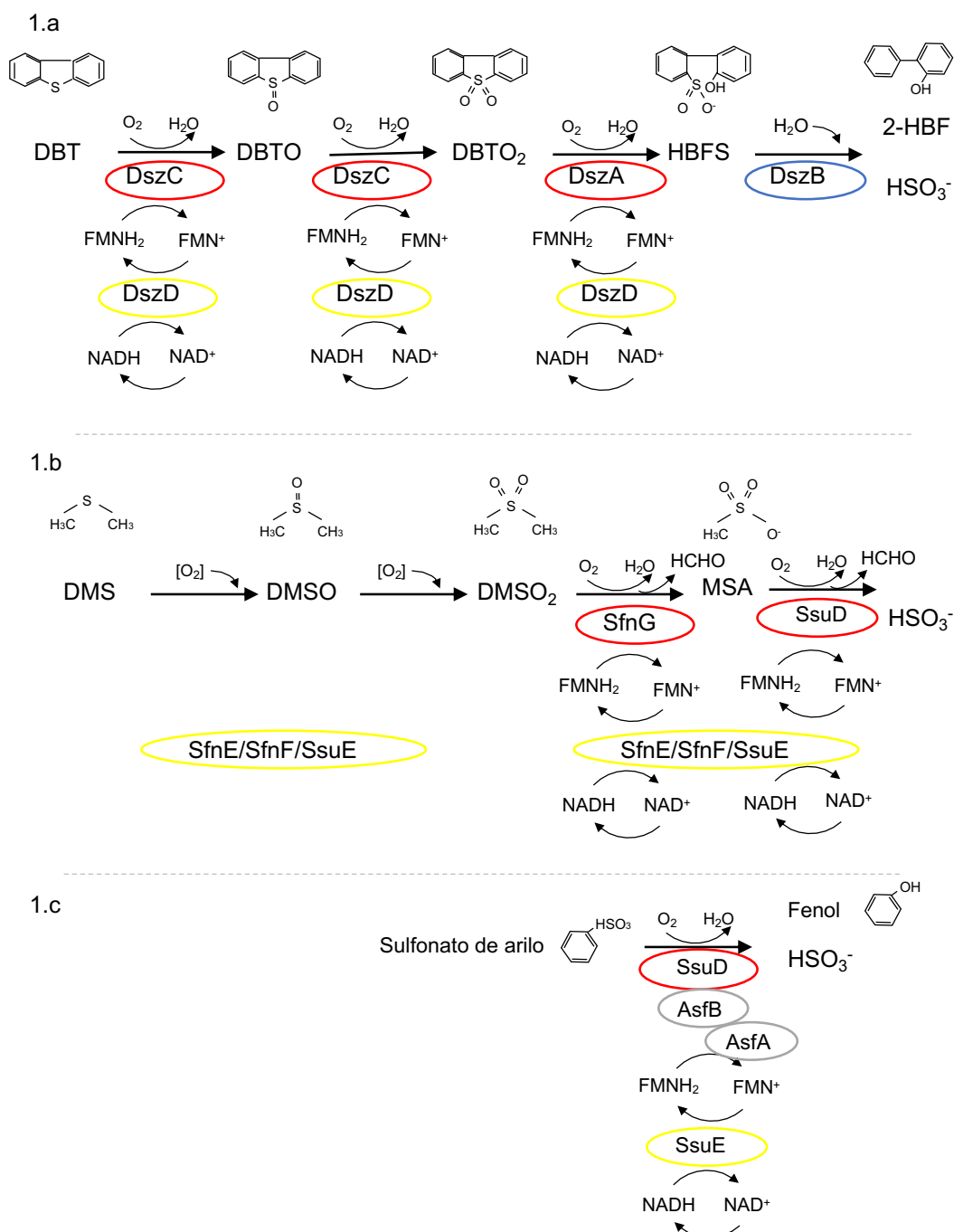


Figura 1: Representación esquemática de rutas de desulfuración de compuestos organosulfurados. **1.a.** Ruta de desulfuración de DBT, adaptada de Martínez *et al.*, 2017. **1.b.** Ruta de desulfuración de DMS/DMSO₂, en *P. putida* DS1 adaptada de Endoh *et al.*, 2003b, Endoh *et al.*, 2005, Kouzuma *et al.*, 2008. **1.c.** Ruta de desulfuración de sulfonato de arilo, en *P. putida* S-313 adaptada de Vermeij *et al.*, 1999, Kertesz *et al.*, 2000. En las tres figuras se rodean en rojo las monooxigenasas, en amarillo los transportadores de electrones, en azul la desulfinasas que cataliza la liberación de azufre y en gris las proteínas con función hipotética.

La desulfuración de DBT, por la vía 4S, que se esquematiza en la figura 1a, es realizada mediante tres monooxigenaciones consecutivas, tras las cuales, se produce una hidrólisis que libera finalmente el bisulfito. Las dos primeras monooxigenaciones están catalizadas por DszC (Dibenzotiofeno monooxigenasa), que transforma el DBT en Dibenzotiofeno 5-óxido ó DBTsulfóxido (DBTO) y posteriormente en Dibenzotiofeno-5,5-dióxido ó DBTsulfona (DBTO₂). A continuación, DszA (Dibenzotiofeno-5,5-dióxido monooxigenasa), incorpora otro átomo de oxígeno, produciendo la escisión oxigenolítica del primer enlace C-S obteniendo como resultado 2-hidroxibifenilo-2-sulfinato (HBFS) el cual es hidrolizado por DszB (2-hidroxibifenilo-2-sulfinato desulfinasa), liberando el ión bisulfito y produciendo 2-hidroxibifenilo (2-HBF), (Gray *et al.*, 1996). DszC y DszA son FMNH₂ dependientes; el flujo de electrones es posible por la acción de la oxidorreductasa DszD, que permite la regeneración de los cofactores necesarios para la acción de DszC y DszA (Kilbane II, 2006).

La acumulación de intermediarios oxidados cuando el petróleo crudo se biodesulfuriza indica que DszA y DszB tienen un rango de sustrato más restringido que DszC, además, se ha demostrado que DszB cataliza el paso limitante en velocidad de la vía de desulfuración para *Rhodococcus erythropolis* IGTS8 (Kilbane II, 2006, Bordoloi *et al.*, 2016). DszB, es la enzima que realmente realiza la desulfuración, siendo este paso el menos comprendido del proceso, ya que los resultados experimentales apuntan a un mecanismo diferente al de otras desulfinasas (Geronimo *et al.*, 2017). Están descritas reacciones de rotura enzimática de enlaces carbono-azufre entre un grupo fenilo y un ácido sulfónico, como en la desulfuración de algunos sulfonatos de arilo; se ha propuesto que esta rotura es catalizada por una monooxigenasa, pero se sabe que DszB no lo es. Por otra parte, también se han descrito roturas de enlaces enzimáticos carbono-azufre entre el carbono alifático y el grupo sulfínico ácido; parece que esa rotura es debida a enzimas que requieren piridoxal 5-fosfato (Gray *et al.*, 2003). Sin embargo, DszB no necesita piridoxal 5-fosfato ni ningún otro cofactor (Geronimo *et al.*, 2017). Geronimo *et al.*, proponen que el mecanismo de acción de DszB, debe ser por sustitución electrofílica por un protón, siendo el residuo C27 el que transferiría su protón al HBFS. De los posibles mecanismos de acción inicialmente propuestos para DszB, este es el de menor energía, y además es consistente con otros estudios de mutagénesis dirigida (Geronimo *et al.*, 2017).

Las monooxigenasas son enzimas que introducen un átomo de oxígeno, en moléculas orgánicas y participan en muchos procesos biológicos. Las monooxigenasas DszA y DszC, son diferentes de otras bien descritas; ya que no poseen un cofactor unido fuertemente, sino que utilizan FMNH₂ como cosustrato (Kertesz, 2000). La familia de

monooxigenasas bacterianas dependientes de FMNH₂ que utilizan FMNH₂ como cosustrato en lugar de un grupo prostético, es relativamente pequeña e incluye las oxigenasas de DBT (DszA y DszC), las sulfonato oxigenasas SsuD y MsuD, las oxigenasas de ácido nitriloacético (NTA), y de ácido etilendiaminotetraacético (EDTA) y una oxigenasa implicada en la síntesis del antibiótico pristamicina IIB. (Kahnert *et al.*, 2000) Las enzimas implicadas en la desulfuración de DBT y en la utilización de diferentes sulfonatos, están muy relacionadas. Al igual que en la desulfuración de DBT, el paso clave en la mayoría de las reacciones de desulfonación es la reacción de oxigenación, en la que la capacidad oxidante del oxígeno se utiliza para desestabilizar o romper el enlace carbono-azufre. La reacción catalizada por DszA recuerda mucho a la escisión monooxigenativa de sulfonatos aromáticos en *P. putida* S-313, que se esquematiza en la figura 1.c, catalizada por SsuD. A pesar de la similitud en el mecanismo de rotura del enlace C-S y que ambas proteínas necesitan FMNH₂ (Vermeij *et al.*, 1999), SsuD de *P. putida* S-313 y DszA sólo tienen un 26% de identidad. SsuD recibe los electrones de SsuE, con la participación de la reductasa putativa AsfA y la ferredoxina putativa AsfB. *P. putida* S-313 es capaz de desulfonar un amplio rango de sulfonatos aromáticos, a partir de los cuales produce el correspondiente fenol y libera azufre (Vermeij *et al.*, 1999) (Kertesz, 2000). A pesar de que inicialmente se pensaba que SsuD de *P. putida*, estaba implicada exclusivamente en la desulfonación de sustratos aromáticos (Kertesz, 2000), posteriormente se ha demostrado que SsuD de *P. putida* DS1 es esencial para el catabolismo de DMS en dicha bacteria (Figura 1.b) (Endoh *et al.*, 2003b).

La vía 4S de desulfuración de DBT presenta similitudes con rutas de degradación de otros compuestos organosulfurados. Es similar a la de desulfuración de DMS, que se esquematiza en la figura 1.b, ya que ambas implican el concurso de 3 monooxigenaciones seguidas. Por otra parte, DszA de *R. erythropolis* IGTS8 y SsuD de *P. putida* DS1, presentan un 31% de identidad (Endoh *et al.*, 2003b). Además, SfnC (monooxigenasa dependiente de FMNH₂) de *P. putida* DS1 y DszC *R. erythropolis* IGTS8 muestran un 43 % de identidad y las monooxigenaciones catalizadas por DszC son similares a las oxidaciones necesarias para convertir DMS a dimetilsulfona (DMSO₂) en la cepa DS1 (Endoh *et al.*, 2003). El patrón de expresión de las proteínas de ambas vías, que únicamente se expresan en limitación de azufre, también es similar. Por lo que se ha especulado con la posibilidad de que las enzimas implicadas en la desulfuración de DMS de *P. putida* DS1 puedan estar relacionadas evolutivamente con las implicadas en la desulfuración de DBT de *Rhodococcus* sp (Endoh *et al.*, 2003b).

De hecho, puesto que, según la base de datos de *P. aeruginosa* PAO1, esta bacteria tiene varios homólogos de *dszC* y *dszA* en su cromosoma, inicialmente se especuló con la posibilidad de que algunos de los homólogos de *dszC* y/o *dszA* del genoma de DS1 pudieran estar involucrados en la oxidación de DMS a DMSO₂ o la conversión de DMSO₂ a MSA (Endoh *et al.*, 2003b). Sin embargo, a pesar de que SfnA y SfnB muestran similitud con DszC, no parecen ser necesarias para el metabolismo de DMS y DMSO (Endoh *et al.*, 2005).

SfnG que muestra un 27% de similitud con DszA, es la monooxigenasa responsable de la conversión de DMSO₂ a MSA. SfnG y DszC pertenecen a la familia de monooxigenasas dependientes de FMNH₂ y es posible que ambas compartan un mecanismo catalítico similar para la escisión del enlace C-S de la organosulfona (Endoh *et al.*, 2005).

Por otra parte, en *Geobacillus thermoglucosidasius* W-2, bacteria termófila capaz de crecer en petróleo pesado y desulfurar varios alcanosulfonatos además de DBT, se ha propuesto que la desulfuración de DBT se llevaría a cabo de una forma “alternativa” a la vía 4S, puesto que convertiría el DBT y otros compuestos heterocíclicos del mismo tipo en sulfonatos de DBT utilizando proteínas similares a DszC y DszA, mientras que la conversión de azufre orgánico en azufre inorgánico se llevaría a cabo por las proteínas codificadas por 2 grupos génicos: *ssuD1* / *ssuE1* y *ssuD2* / *ssuE2*, rompiendo en enlace carbono-azufre de forma similar a como lo hace DszB. (Peng *et al.*, 2019). Sin embargo, a pesar de la aparente similitud en los mecanismos de la reacción de escisión del enlace C-S, el sistema Ssu y el sistema Dsz son distintos catalíticamente; *P. putida* S-313 no puede desulfurar DBT (Vermeij *et al.*, 1999) y una cepa de *R. erythropolis* que fue capaz de crecer con sulfonatos aromáticos y que contenía el gen *ssuD* no creció con DBT (Kertesz, 2000).

Además, en *Corynebacterium glutamicum* ATCC 13032, que es capaz de crecer con varios sulfonatos y ésteres de sulfonato como fuentes de azufre, se han encontrado genes *ssu* que codifican para proteínas muy parecidas a las Ssu de *Escherichia coli* o de *Bacillus subtilis* (*ssuC*, *ssuB* y *ssuA*, son necesarios para la utilización de sulfonatos alifáticos) y genes *seu*, que codifican para enzimas que tienen similitud con las enzimas Dsz de *Rhodococcus* IGTS8, y se ha comprobado que los genes *seuA*, *seuB* y *seuC* pueden reemplazar a *ssuD1* y *ssuD2* (Koch *et al.*, 2005).

La vía 4S de desulfuración de DBT requiere poder reductor en forma de FMNH₂. En relación con el aporte de electrones necesarios para llevar a cabo el proceso, merece especial atención DszD. Se trata de una FMN reductasa dependiente de NADH responsable de proporcionar los electrones necesarios tanto a DszA como a DszC, lo

cual, se ha podido determinar tanto con experimentos *in vitro* como *in vivo* (Gray *et al.*, 2003). Sin embargo, a pesar de su importancia en el proceso, está muy poco conservada y no parece tener otra función fisiológica adicional, ya que mutaciones por inserción en el gen *dszD* de *R. erythropolis* producen la pérdida tanto de la actividad de la flavina reductasa como del fenotipo de desulfuración, pero, sin embargo, no tienen ningún efecto aparente sobre la viabilidad de la cepa cuando se cultiva en otra fuente de azufre (Gray *et al.*, 2003).

Las flavinas reductasas son enzimas definidas por su capacidad para catalizar la reducción de flavinas libres, riboflavina, FMN o FAD, mediante piridín nucleótidos reducidos. El papel de la enzima sería proporcionar un sitio activo que acomoda transitoriamente al piridín nucleótido reducido y la flavina, favoreciendo la transferencia directa de hidruro (Gray *et al.*, 2003). Dado que DszD utiliza FMN como sustrato disociable, se ha clasificado como flavina reductasa de clase I. Esta clase se subdivide en base a similitud de secuencia, perteneciendo DszD al mismo subgrupo que ActVB, SnaC y VImR, que participan en la biosíntesis de los antibióticos actinorodina en *Streptomyces coelicolor*, pristamicina en *Streptomyces pristinaespiralis*, y valanimicina en *Streptomyces viridifaciens*, HpaC implicada en la oxidación del 4-hidroxifenilacetato en *E. coli* y cB en la degradación de nitriloacetato en *Chelatobacter heintzii* (Filisetti *et al.*, 2002). Los datos sugieren que el mecanismo enzimático de las proteínas ActVB y DszD es secuencial, uniéndose primero el NADH. La flavina libre reducida es muy inestable y se autooxida con gran rapidez; por lo tanto, se supone que se produce algún tipo de interacción proteína-proteína entre la reductasa y la oxigenasa para estabilizar la flavina reducida y permitir la inserción del oxígeno (Gray *et al.*, 2003).

Los genes *dszA*, *dszB*, *dszC* de *Rhodococcus* IGTS8 forman parte de un plásmido, de 120 kb. El fenotipo de desulfuración es conferido por un fragmento de 4 kb de dicho plásmido. *dszA*, *dszB*, *dszC* se transcriben en un operón en la misma dirección y bajo el control de un solo promotor; sin embargo, el gen que codifica para DszD, se encuentra en el cromosoma (Mohebbi & Ball, 2008), de hecho, en ninguna bacteria degradadora de DBT por la vía 4S caracterizada hasta el momento, aparece la reductasa unida al operón *dsz*. Ishii *et al.*, (Ishii *et al.*, 2000b), clonaron el gen *tdsD*, que codifica la flavina reductasa de la ruta de DBT de la bacteria termófila, *Paenibacillus* sp. A11-2 (TdsD); comprobando que *tdsD* se localizaba a cierta distancia del operón formado por *tdsA*, *tdsB* y *tdsC* (Ishii *et al.*, 2000b) Así mismo, Furuya *et al.*, (Furuya *et al.*, 2005), clonaron los genes de desulfuración de DBT *bdsABC* de *M. phlei* WU-F1, comprobando que no existía ningún gen de flavina reductasa cerca de los genes *bdsABC*, al menos en 3,3 kb aguas arriba y 0,4 kb aguas abajo. Posteriormente, clonaron el gen (*frm*) que codifica la flavina reductasa de la ruta. Aunque las secuencias

de aminoácidos de *bdsABC* muestran una identidad del 70% con los de los genes *dszABC* *Rhodococcus* IGTS8, *frm* solo tiene un 33% de identidad con *dszD* de dicha bacteria. Lo que sugieren que los genes de desulfuración de DBT y el gen de la flavina reductasa han evolucionado de forma independiente (Furuya *et al.*, 2005).

De hecho, las monooxigenasas de la ruta de desulfuración de DBT no parecen ser restrictivas en cuanto a la reductasa de la que reciben los electrones, ya que, los genes *dszABC* (o similares) de distintas especies se han subclonado en diferentes estirpes, y han presentado actividad degradadora de DBT, habiendo sido suplida la actividad de la reductasa, con otra de la bacteria hospedadora, como demuestra por ejemplo el caso de Gallardo *et al.*, en el que subclonaron exclusivamente el operón *dszABC* de *Rhodococcus erythropolis* IGTS8, integrándolo en *Pseudomonas* y consiguiendo actividad desulfuradora de DBT (Gallardo *et al.*, 1997). Además, se ha demostrado que la actividad de DszD puede ser reemplazada *in vitro* por otras FMN reductasas (Kertesz, 2000) (Gray *et al.*, 2003). Galán *et al.*, 2000, demostraron que HpaC, una oxidorreductasa de *E. coli* W, puede suministrar los electrones de su FMNH₂ a las monooxigenasas Dsz. Además, mediante la sobreexpresión del gen *hpaC* en cepas de *Pseudomonas* que llevan el grupo de genes *dszABC* se consiguió una alta tasa de eliminación de azufre *in vivo* (Galán *et al.*, 2000), lo que sugiere que el suministro de flavina reducida a la monooxigenasa es un paso limitante de la velocidad en la vía *in vivo* completa (Gray *et al.*, 2003).

P. putida DS1 posee varias reductasas susceptibles de proporcionar FMNH₂ para las monooxigenasas. SfnF, que es una putativa FMN reductasa dependiente de NADH, SsuE, FMN reductasa supuestamente dependiente de NADPH esencial para la utilización de DMSO₂ (Endoh *et al.*, 2005), y SsuE, cuya función se ha comparado con la de DszD (Endoh *et al.*, 2003b). A pesar de que, como ya se ha comentado, las enzimas de la ruta de DBT y las de utilización de sulfonatos están muy relacionadas, no ocurre lo mismo con la organización de los genes que codifican para dichas enzimas, ya que, en *P. putida* DS1, *ssuE*, y *ssuD*, si aparecen codificadas en el mismo operón, *ssuEADCBF* (Endoh *et al.*, 2003b). SsuD presenta homología con DszA, aunque su actividad catalítica es diferente, puesto que SsuD cataliza la liberación del ión bisulfito. Es bastante sorprendente que el operón de los genes de degradación de DBT no incluya el transportador de electrones, como sucede en las rutas de degradación de contaminantes orgánicos, en las que los genes que codifican para los transportadores de electrones a las monooxigenasas o dioxigenasas van ligados a los genes codificantes de las oxigenas.

Los primeros genes responsables de la degradación de DBT por la vía 4S caracterizados, fueron los de *Rhodococcus erythropolis* IGTS8, inicialmente nombrados como *soxABC* (Denome et al., 1994), y posteriormente denominados *dszABC* por Piddington et al., (Piddington et al., 1995). Posteriormente se han descrito genes *dsz* (o similares) en gran variedad de microorganismos, con un alto grado de conservación entre las distintas cepas con actividad desulfuradora de DBT. Se han aislado del suelo un gran número de microorganismos con capacidad desulfuradora cuyos genes *dsz* presentan secuencias casi idénticas (identidad en general en torno al 90%) a los genes de *R. erythropolis* IGTS8, como los genes *dsz* de *Gordonia alkinovorans* 1B, que muestran aproximadamente un 88% de identidad con los de *R. erythropolis* y los genes de desulfuración de la bacteria termófila *Paenibacillus* sp. A11-2 (*tdsABC*) que muestra una homología del 73%, 61% y 52% respectivamente con los genes *dszABC* (Kilbane II, 2006). Esta similitud, unida al hecho de que en algunos casos estos genes estén asociados a elementos móviles (Denome et al., 1994, Ishii et al., 2000a) apoya la hipótesis de que estos genes están comúnmente sujetos a transferencia horizontal en la naturaleza (Mohebbi & Ball, 2008).

La organización génica de los genes *dszABC* está muy conservada. Los genes se organizan en un bloque muy compacto, con muy poca separación entre *dszB* y *dszC* (por ejemplo, de 13 pb en *R. erythropolis* IGTS8) e incluso solapamiento entre *dszA* y *dszB*, aunque la longitud del solapamiento si varía de unas bacterias a otras desde 4 pb en *R. erythropolis* IGTS8 hasta de 64 pb en *G. alkinovorans* 1B, (Kilbane II, 2006). Las secuencias Shine-Dalgarno son perfectamente identificables aguas arriba de los tres codones de inicio (Li et al., 2007). La organización génica del operón parece tener un efecto directo en la eficiencia del proceso de BDS: con ensayos de RT-qPCR se ha comprobado que la transcripción de los genes *dsz* en *Rhodococcus erythropolis* es decreciente desde *dszA* hasta *dszC*; siendo la cantidad de ARNm de *dszB* sólo un 30% que la de *dszA*. Por otra parte, mediante ensayos de *western blot*, se comprobó que la expresión de DszB es mucho menor que la de DszC. Ambos resultados sugieren que la traducción de *dszB* es menos eficiente que la de *dszA* y *dszC*, sugiriendo que dicha traducción, se reprime por una estructura situada antes de su codón de inicio, el cual solapa con el codón de parada de *dszA*. Li et al, rediseñaron el operón *dsz*, eliminando el solapamiento entre *dszA* y *dszB*, con lo que consiguieron mejorar la actividad de desulfuración de DBT de *R. erythropolis* DR-1 aproximadamente cinco veces con respecto a la estirpe original (Li et al., 2007, Li et al., 2008).

La aplicación directa de la BDS en la eliminación del azufre de los crudos de petróleo tiene dos problemas técnicos que dificultan su utilización en la industria: en

primer lugar, la dificultad de desarrollar un bioproceso continuo, y en segundo lugar los problemas para desarrollar biocatalizadores robustos con altas tasas catalíticas de desulfuración. Estos problemas se derivan fundamentalmente del hecho de que los genes *dsz*, por regla general, se expresan a niveles bajos en sus hospedadores naturales, ya que las cantidades de azufre necesarias para el crecimiento bacteriano son bajas, y la regulación de los genes *dsz*, en muchos casos depende de azufre, lo que conlleva que las tasas de desulfuración de DBT sean bajas para las necesidades de la industria. (Martínez *et al.*, 2016). La regulación global por azufre, así como la regulación específica de los genes *dsz*, se abordarán más adelante.

Por otra parte, el principal cuello de botella de la ruta de biodesulfuración es la inhibición de la vía causada por la propia liberación del producto 2-HBF (Abin-Fuentes *et al.*, 2013) (Kilbane II, 2016), siendo además DszC, la enzima iniciadora del proceso, la que sufre esta inhibición de forma más severa. DszC muestra un patrón de inhibición típico de “inhibición por retroalimentación” no sólo por 2-HBF, sino también por HBFS, el último intermediario de la vía, (Abin-Fuentes *et al.*, 2013) (Li *et al.*, 2019). En *R. erythropolis* IGTS8 se ha descrito que la IC_{50} de DszA, DszB y DszC con respecto a 2-HBF es de 60, 110 y 50 μM , respectivamente, y la IC_{50} de DszC por HBFS es de 15 ± 2 μM . Estas cuatro interacciones inhibitoras, son muy probablemente las responsables de la reducción en la actividad del biocatalizador que se observa durante un proceso típico de BDS. La inhibición causada por HBFS sobre DszC, mantiene una baja tasa de BDS al inicio del proceso, aunque a medida que este producto se va metabolizando, aumentan los niveles de 2-HBF, el cual, al acumularse, causa inhibición sobre DszA, DszB y DszC. Por otra parte, la eficiencia de la actividad catalítica (o constante de especificidad, definida como la relación k_{cat}/K_m) de las enzimas de *R. erythropolis* IGTS8 en relación con sus sustratos, es de: 3.1, 1.3, 1.1 $\mu M^{-1} min^{-1}$, para DszA, DszB y DszC, respectivamente, y la eficiencia catalítica de DszD para NADH y FMN es de 6.7 y 100 $\mu M^{-1} min^{-1}$ respectivamente; con lo que la eficiencia relativa de las enzimas entre sí es $DszD > DszA > DszB \approx DszC$. Considerando la baja eficiencia catalítica de DszC, así como su sensibilidad tanto a 2-HBF como a HBFS, y siendo además la enzima iniciadora del proceso, se puede considerar como la limitante para la BDS. (Abin-Fuentes *et al.*, 2013). (Li *et al.*, 2019). La inhibición de 2-HBF sobre DszC es una inhibición no competitiva, ya que el 2-HBF muestra afinidad tanto por la enzima como por el complejo enzima-sustrato, y la velocidad de la actividad enzimática de DszC se reduce, sin cambiar la K_m a medida que se aumenta la concentración de 2-HBF (Abin-Fuentes *et al.*, 2013). Además, aunque los sitios de unión de HBF no están bien caracterizados, se sabe que son diferentes a los de DBT. (Li *et al.*, 2019).

Se han abordado diferentes estrategias, para conseguir una BDS más eficiente, estas estrategias se pueden clasificar en mejoras biológicas y mejoras del proceso. Las principales se relacionan a continuación.

MEJORAS BIOLÓGICAS

-Empleo de promotores heterólogos. El operón *dszABC* de *R. erythropolis* /GTS8 está fuertemente reprimido por sulfato y aminoácidos que contienen azufre, lo que constituye un problema importante para diseñar biocatalizadores eficientes. Para aliviar esta represión, el operón *dsz* se clonó bajo el control de señales reguladoras heterólogas de amplio espectro (*lacIq* / *P_{trc}*) y se introdujo en el cromosoma de diferentes cepas de *Pseudomonas* consiguiendo desulfurar DBT de manera eficiente en medios de cultivo que contenían azufre. (Gallardo et al., 1997).

-Reordenamiento de los genes *dsz*. Como ya se ha comentado, En *Rhodococcus erythropolis*, se ha podido comprobar que los niveles de transcripción y traducción de *dszA*, *dszB* y *dszC* disminuyen según las posiciones de los genes en el operón *dsz*, y que al reorganizar los genes *dszABC* en diferente sentido (*dszCBA*), la relación de los ARN mensajeros de *dszA*, *dszB* y *dszC*, varió de: 11:3,3:1 a 1:16:5, con ensayos de western blot se comprobó que aumentaron los niveles de expresión de *dszB* y *dszC*, aunque disminuyeron los de *dszA*, consiguiendo aumentar la capacidad de desulfuración 12 veces con respecto al silvestre. El reordenamiento implicó eliminar el solapamiento entre *dszA* y *dszB* y eliminar la estructura que reprime la traducción de *dszB*. A su vez, con esta configuración no se acumula HBFs, lo cual tiene un efecto directo en la mejora de la BDS (Li et al., 2007, Li et al., 2008).

-Evolución dirigida de las enzimas *Dsz*. Aplicada, por ejemplo, para obtener versiones más eficientes de *DszC*. Recientemente, Li et al., han obtenido una versión de *DszC* insensible a 2-HBF mediante mutagénesis, de saturación iterativa. Al sobreexpresar dicha proteína, se mejoró sustancialmente la tasa de conversión específica (Li et al., 2019). La evolución dirigida se ha realizado también incorporando al operón *dsz* un gen que codifica para un polipéptido rico en metionina y cisteína, aminoácidos que contienen azufre, de forma que, al expresar los genes *dsz* para metabolizar DBT la bacteria tendría que expresar dicho polipéptido, lo cual aumentaría las necesidades de azufre de la célula, y se podrían seleccionar mutantes con mayor expresión del operón *dsz*, ya que la bacteria requeriría permanentemente más azufre. El gen se insertó entre *dszA* y *dszB*, con lo que, además, se elimina el solapamiento. Al introducir este operón sintético en una estirpe incapaz de crecer en DBT como única fuente de azufre y realizar varios pases en medio mínimo con DBT como única fuente

de azufre, se seleccionaron candidatos cuya capacidad de desulfuración superaba en 20 veces la de *Rhodococcus erythropolis* IGTS8 (Pan *et al.*, 2013).

-Modificaciones en la membrana de la célula. Es un objetivo importante puesto que la membrana debe jugar un papel en el trasiego de metabolitos durante la BDS. Parece que al adicionar DBT al medio, éste entra en la célula por difusión pasiva. Dentro del citoplasma, se transforma en 2-HBF que, al tratar de salir, queda atrapado en parte en la membrana pudiendo, por tanto, ser responsable de la retención de 2-HBF dentro de los biocatalizadores. (Abin-Fuentes *et al.*, 2013). Como resultado de una mutagénesis por transposición, Watanabe *et al.*, aislaron una estirpe de *Rhodococcus erythropolis* cuya composición en ácidos grasos de membrana, era diferente a la estirpe original con una relación de ácidos grasos saturados/insaturados de 13 en la estirpe mutante vs. 2,3 en la silvestre, lo cual aumentaba la fluidez de su membrana, mejorando la transferencia de DBT y sus derivados con grupos alquilo. La nueva estirpe era capaz de consumir 50 mg/l azufre presente en el gasóleo ligero en 2 h, mientras que la estirpe parental era incapaz de desulfurarlo (Watanabe *et al.*, 2003).

-Empleo de estirpes tolerantes a solventes orgánicos. Utilizadas en sistemas bifásicos. Dichos sistemas, que son útiles porque aumentan la productividad de la biocatálisis, contienen disolventes orgánicos inmiscibles en agua, muchos de los cuales son muy tóxicos y pueden ser letales para la mayoría de los microorganismos incluso en concentraciones bajas. (Tao *et al.*, 2006, Tao *et al.*, 2011)

-Empleo de estirpes capaces de producir biosurfactantes. Los biosurfactantes aumentan el contacto entre el aceite y la fase acuosa aumentando la concentración de los compuestos hidrófobos en ella, y por tanto la biodisponibilidad de los compuestos tiofénicos, mejorando la transferencia de masa, y evitando la acumulación de 2-HBF. (Gallardo *et al.*, 1997) (Nassar *et al.*, 2020).

-Empleo de estirpes capaces de formar biofilm. Los biofilms son más resistentes a agresiones físicoquímicas que los cultivos celulares, pudiendo resistir mejor las duras condiciones del biocatalizador. (Martínez *et al.*, 2017). Recientemente se ha construido una estirpe de *R. erythropolis* superproductora de biofilm, gracias a la expresión heteróloga del gen *adrA*, que codifica para una diguanilato ciclasa, que promueve la síntesis de di-GMP-cíclico, molécula desencadenante de la formación de biofilm. La nueva estirpe, mejora la capacidad biodesulfuradora de la estirpe original (Dorado-Morales *et al.*, 2021).

-Empleo de estirpes y enzimas termófilas. Debido a que el tratamiento de refinado del petróleo, así como la HDS (en el caso de combinarla con BDS), necesitan altas temperaturas, puede ser útil emplear estirpes y/o enzimas que puedan resistir dichas temperaturas, para evitar tener que introducir pasos de enfriamiento en el

proceso. (Martínez et al., 2017). El gen *dszC* de *R. erythropolis* IGTS8 se ha podido expresar en *T. thermophilus* (Park et al., 2004), que puede crecer a temperaturas entre 55 °C y 85 °C, y es probable que pueda expresarse la vía de desulfuración completa en dicho organismo. Además, podría emplearse evolución dirigida y presión selectiva para desarrollar enzimas de desulfuración capaces de funcionar a temperaturas más altas. (Kilbane II, 2006). Aunque, por supuesto, también sería interesante seguir insistiendo en el aislamiento de bacterias termófilas capaces de crecer con DBT como única fuente de azufre a alta temperatura.

-Empleo de sistemas de transporte para mejorar el flujo de compuestos de baja solubilidad. Para facilitar la entrada de DBT en la célula, se ha empleado un sistema de transporte de compuestos hidrofóbicos codificado por *hucABC* de *P. aeruginosa* NCIMB9571 (Noda et al., 2003), el cual se clonó en *P. putida* IFO13696, que contenía un operón artificial *dszABCD* de *R. erythropolis* KA2-5-1 inserto en su cromosoma y que puede desulfurar DBT disuelto en agua, pero no en N-tetradecano, consiguiendo así que la nueva estirpe fuera capaz de desulfurar DBT disuelto en N-tetradecano, (Noda et al., 2003b).

-Desarrollo de nuevas estirpes que reúnan diferentes estrategias de mejora. Se trata de reunir distintas características de interés junto con la capacidad de desulfuración, por ejemplo, la tolerancia a solventes orgánicos. Tao et al., introdujeron los genes de biodesulfuración de *Rhodococcus erythropolis* XP en *P. putida* Idaho, bacteria tolerante a solventes orgánicos, obteniendo *Pseudomonas putida* A4, la cual fue capaz de crecer en cultivo bifásico con varios compuestos heterocíclicos de azufre, y en particular siendo capaz de degradar DBT, en presencia de varios solventes orgánicos. (Tao et al., 2006). Con el mismo objetivo, y puesto que *Pseudomonas putida* es un buen candidato para el desarrollo de biocatalizadores, desarrollaron una nueva estirpe, *P. putida* DS23, tolerante a solventes orgánicos y capaz de desulfurar DBT. Para lo cual previamente construyeron un vector con los genes de desulfuración de DBT bajo el control de P_{srp} , en lugar de P_{tac} . P_{srp} controla la expresión de una bomba de eflujo en *P. putida* S12, y que puede ser inducido por disolventes aromáticos y alifáticos. La nueva estirpe, mostró una alta eficiencia de degradación de DBT en ensayos de células en reposo en medio bifásico (Tao et al., 2011).

-Desarrollo de casetes génicos. Para poder expresarlos en diferentes estirpes que reúnen diferentes ventajas. *Pseudomonas* es un buen candidato para ser empleado en BDS debido a que es uno de los microorganismos mejor estudiados y más abundantes en los crudos, además de que ciertas estirpes son capaces de producir biosurfactantes, tienen alta tolerancia a solventes y alta resistencia a metales pesados, también presentes en combustibles fósiles. (Gallardo et al., 1997). Gallardo et al.,

desarrollaron un casete con los genes de *Rhodococcus erythropolis* IGTS8, bajo el control de un promotor heterólogo, que se introdujo en el cromosoma de *Pseudomonas putida*; obteniendo una estirpe capaz de desulfurar DBT y de producir biosurfactantes. Basándose también en los genes *dsz* de *R. erythropolis* IGTS8, y empleando como hospedador *P. putida* KT2440, Martínez *et al.*, trataron de abordar varias de las estrategias de mejora ya comentadas: reordenamiento de genes, y eliminación del solapamiento, para mejorar la expresión de *dszB* (que codifica el paso que limita la velocidad el proceso de BDS), así como la expresión de los genes bajo promotores heterólogos, mejora de la traducción (mediante el empleo de RBS de *P. putida* y del uso mejorado de codones), sustituciones de aminoácidos para hacer DszB, más eficiente y sobre todo la incorporación de DszD de *P. putida*, que no forma parte de la ruta original de *R. erythropolis*. Además, un paso adicional en la implementación de los casetes génicos como estrategia de mejora, fue la compartimentalización de los diferentes pasos de la vía 4S en distintos casetes y la combinación de esta estrategia con el desarrollo de consorcios, de forma que se puedan modificar las proporciones de cada parte del proceso, para conseguir mayores tasas de BDS. Puesto que, como se ha comentado, un punto crítico es la inhibición causada por 2-HBF y HBFS es muy importante controlar tanto los niveles de DszD, que tendrán un efecto directo en el funcionamiento de DszC y DszA, así como los niveles de DszB, y, sobre todo, los niveles relativos de las enzimas entre sí, para hacer el proceso más eficiente. (Martínez *et al.*, 2016).

-Empleo de la herramienta de edición génica CRISPR-Cas, para desarrollar organismos con una BDS más eficiente. Tal y como argumentan Parveen *et al*, 2020, la herramienta CRISPR-Cas supone una ventaja frente a métodos convencionales de ingeniería genética, puesto que presenta alta especificidad, eficiencia y robustez, así como facilidad de manejo, por lo que proponen un modelo para desarrollar bacterias desulfuradoras insertando modificaciones genéticas en su cromosoma. Según el modelo propuesto, se podría introducir un operón *dsz* en el genoma de ciertos hospedadores como *E. coli*, *Ralstonia*, *Rhodococcus*, *Corynebacterium*, *Paenibacillus* o *Pseudomonas*, bajo el control de un promotor fuerte no reprimible por sulfato, incluyendo además el gen para la oxidorreductasa bajo el control del mismo promotor, de forma más rápida y menos laboriosa que por métodos convencionales (Parveen *et al*, 2020).

-Empleo de estirpes capaces de realizar la “vía 4S extendida o desulfuración profunda”. Como se ha mencionado anteriormente, algunas bacterias desulfuradoras, metabolizan el 2-HBF, para dar 2-metoxibifenilo (2-MBF), lo cual, supone una ventaja para la posible utilización industrial de la BDS, ya que elimina la inhibición causada por 2-HBF. Además, el 2-metoxibifenilo es en general, menos tóxico para la célula que el 2-HBF. Hay pocos ejemplos de microorganismos descritos capaces

de realizar este proceso, entre ellos *Mycobacterium goodii* X7B bacteria termófila, que también puede desulfurar gasolina (Li *et al.*, 2005), *Achromobacter sp.* (Bordoloi *et al.*, 2014), capaz de utilizar DBT y 4-metilDBT como única fuente de azufre, consiguiendo desulfurar el 7% de azufre del gasóleo, tratando 5 ml de gasóleo con 14 g de peso seco, y *Chelatococcus sp.* capaz de desulfurar DBT, 4-metilDBT y 4,6-dimetilDBT, pudiendo reducir el contenido de azufre del diésel en un 12%, tras 24 horas de tratamiento con 1 gramo de células en reposo. (Bordoloi *et al.*, 2016).

MEJORAS DEL PROCESO

-Empleo de células en reposo, aunque la utilización de cultivos bacterianos es útil para verificar la capacidad de diferentes microorganismos para realizar BDS, el uso de células en reposo mejora la eficiencia de la BDS y permite desulfurar mayores cantidades de DBT (Martínez *et al.*, 2017).

-Optimización del proceso mediante métodos estadísticos. Con la aplicación de la Metodología de superficie de respuesta (RSM), que se utiliza para encontrar una relación entre ciertas variables efectivas y una o más respuestas en un sistema, Derikvand *et al.*, definieron como variables de operación a tener en cuenta en cultivo acuoso y bifásico en la estirpe *Paenibacillus validus* PD2, los siguientes: (i) la concentración inicial de DBT, ya que *a priori* es tóxico y altas concentraciones podrían causar una disminución de la BDS (aunque debido a su baja solubilidad, también es probable que su efecto tóxico sobre la célula sea muy limitado); (ii) la temperatura, puesto que DszC y DszB parecen ser muy sensibles a la temperatura, y esto se considera un factor limitante de la BDS; (iii) el pH, por su efecto sobre la actividad enzimática. Concluyeron que la eficiencia de BDS en el sistema bifásico es relativamente más alta que la del sistema acuoso, siendo la concentración óptima de DBT en el sistema acuoso de 0,41mM, mientras que en el sistema bifásico es de 7,86 mM. Por otra parte, la temperatura y el pH óptimos son iguales en los dos sistemas (temperatura mesofílica y pH cercano al neutro). (Derikvand *et al.*, 2015).

Así mismo, Nassar *et al.*, realizaron una optimización estadística basada en la técnica One-Factor-At-A-Time (OFAT) y la metodología de superficie de respuesta (RSM), para determinar las condiciones idóneas de desulfuración de la bacteria Gram positiva *Paenibacillus glucanolyticus* HN4, tanto en cultivo bifásico como con células en reposo en un reactor bifásico. Consiguiendo desulfurar el 80% del azufre del crudo. Esta bacteria reúne algunas de las características ya mencionadas como interesantes para una BDS eficiente, como la tolerancia a hidrocarburos, la producción de biosurfactantes y la formación de bioemulsiones. (Nassar *et al.*, 2020).

-Empleo de microorganismos inmovilizados en nanopartículas magnéticas. La inmovilización puede mejorar la estabilidad, prolongando su vida útil, y facilitar la reutilización de los biocatalizadores. (Li *et al.*, 2005), ya que la separación de la mezcla microorganismos-aceite-agua es un grave problema y aún no se ha implementado un procedimiento de separación satisfactorio y económico (Martínez *et al.*, 2017).

-Utilización de adsorbentes selectivos de 2-HBF. La eliminación continua de 2-HBF puede reducir su acumulación en el interior celular y mitigar la pérdida de actividad de desulfuración en el biocatalizador (Abin-Fuentes *et al.*, 2013).

Aparte de la desulfuración del petróleo para combustibles, los biocatalizadores de desulfuración pueden tener muchas otras aplicaciones, que pueden ser viables a menor escala en procesos comerciales, como la producción de compuestos químicos de mayor valor añadido a partir de los diferentes compuestos susceptibles de sufrir desulfuración por la vía 4S, como sulfuros, disulfuros, mercaptanos, sulfóxidos, sulfonas, sulfinatos, sulfonatos, tiofenos y benzotiofeno. (Kilbane II, 2016) (Martínez *et al.*, 2017).

Algunas de estas aplicaciones pueden ser: Descontaminación de zonas expuestas a agentes de guerra química (Kilbane II, 2016). Producción de HBFS que es de interés para la síntesis química de biosurfactantes (Martínez *et al.*, 2016). Producción de 2-HBF que puede usarse como precursor de otros compuestos fenólicos para la síntesis de polímeros, o para fungicidas, desinfectantes, conservantes, colorantes o antioxidantes. Algunas de las enzimas de desulfuración también podrían usarse para producir nuevos monómeros necesarios para producir politioésteres empleados como termoplásticos y películas poliméricas biodegradables. (Martínez *et al.*, 2017). Algunas monooxigenasas de DBT, pueden transformar el indol en índigo, lo cual puede emplearse en búsquedas de nuevos genes que codifiquen para dioxigenasas o para monooxigenasas hidroxilantes de anillo, que puedan tener nuevas actividades de interés (Terrón *et al.*, 2016) (Martínez *et al.*, 2017). Producción de medicamentos, entre los cuales se pueden destacar la producción de nuevos antibióticos, de especial interés debido al aumento de la resistencia de los microorganismos a los antibióticos existentes, así como la producción de otros medicamentos como antitumorales (calicheamicina y bleomicina) y tiazolidinedionas para el tratamiento de la diabetes. (Kilbane II, 2016).

II. REGULACIÓN

II.1. REGULACIÓN DE LOS GENES *dsz*

Aunque se conoce muy poco del mecanismo de regulación de los genes de degradación de DBT, lo que sí parece claro tanto en *Rhodococcus* (Li *et al.*, 1996), como en *Gordonia* (Alves *et al.*, 2007; Feng *et al.*, 2016) y en *Mycobacterium* (Takada *et al.*, 2005) es que los genes *dsz* están fuertemente reprimidos en presencia de fuentes preferenciales de azufre (sulfato, metionina, cisteína, ácido metanosulfónico y casaminoácidos) (Mohebbi & Ball, 2008) y desreprimidos en presencia de DBT o DMSO (Kertesz, 2000). La actividad de las proteínas implicadas en la desulfuración de DBT no se ve inhibida por sulfato, si no que la represión causada por sulfato es a nivel de síntesis de dichas proteínas, observándose una fuerte represión por sulfato de la transcripción desde el promotor de los genes *dsz* (Li *et al.*, 1996, Mohebbi & Ball, 2008). Algunos estudios apuntan a un control negativo de los genes *dsz*, ya que mutaciones aguas arriba del promotor, aumentan la expresión y reducen la represión en *Rhodococcus erythropolis* IGTS8, sugiriendo que, si hay una proteína represora, se une en un sitio que solapa con el promotor. (Li *et al.*, 1996), (Martínez *et al.*, 2017).

Darzins & Mrachko, definieron la secuencia 400pb aguas arriba del codón de inicio de *dszA*, como secuencia con “elementos reguladores, o elementos promotores”, que controlan la transcripción del operón *dsz* en *Sphingomonas* sp. AD109. Sin embargo, esta región no mostró similitud con el promotor de *R. erythropolis* IGTS8 (Darzins & Mrachko, 2000).

Se ha descrito que mutaciones puntuales en la cistationina β -sintasa, de *Rhodococcus erythropolis* KA2-5-1, desreprimen la actividad de los genes *dsz* en presencia de sulfato. Por lo tanto, parece que para que se produzca represión por sulfato en esta estirpe, es necesario que haya asimilación del azufre (Tanaka *et al.*, 2002). La desrepresión de los genes *dsz*, por tanto, no sería una respuesta específica al DBT, sino que sería más bien una respuesta a la limitación de azufre.

Chelatococcus sp, es una bacteria Gram negativa, que se aisló de una muestra de suelo contaminado con crudo de petróleo y que, como se ha comentado, puede realizar la “vía 4S extendida”, pudiendo desulfurar DBT, 4-metilDBT y 4,6-dimetilDBT. El análisis del perfil proteómico de *Chelatococcus* sp. crecido en DBT como fuente de azufre, permitió identificar una gran variedad de enzimas con diversas funciones, entre las que destacan las implicadas en el proceso de desulfuración (se han identificado DszC, DszA, DszB y DszD), proteínas implicadas en la asimilación de azufre y proteínas que pueden estar implicadas en el transporte metabolitos tóxicos como 2-HBF o 2-MBF,

y sulfito (como una enolasa y una ATP sintasa), así como proteínas de la familia de transportadores MFS (“Major Facilitator Superfamily”), que pueden transportar DBT y sus metabolitos a través de la membrana celular (Bordoloi *et al.*, 2016). Es interesante también tener en cuenta que el DBT y los productos finales de su degradación pueden suponer una fuente de stress para la célula, habiéndose podido comprobar en este estudio, que *Chelatococcus sp.* crecido en DBT como fuente de azufre, regula positivamente proteínas de respuesta a estrés, como HSP-10 y HSP-60, y antioxidantes (Bordoloi *et al.*, 2016).

En cuanto a proteínas reguladoras implicadas en el control de la ruta de desulfuración de DBT, hay pocos ejemplos descritos. En *Gordonia sp.* IITR100, bacteria Gram positiva perteneciente a *Actinobacteria* y conocida por su capacidad de metabolizar compuestos orgánicos de azufre, se ha caracterizado un regulador transcripcional de la familia TetR que se une a secuencias aguas arriba del promotor de los genes *dszABC* activando la expresión del operón, comprobándose además que este activador también activa la expresión de los genes *dsz* en otros hospedadores como *Rhodococcus erythropolis* y *E. coli* (Murarka *et al.*, 2019a, Murarka & Srivastava, 2019b). Por otra parte, en un ensayo “pull-down” con células de *Gordonia sp.* IITR100 crecidas en sulfato, se aisló la proteína WhiB1 capaz de unirse al promotor de los genes *dsz*. Es una proteína de la familia WhiB de reguladores transcripcionales, únicamente descritos en *Actinobacterias*, que se unen a ADN y pueden estar involucrados en diversas funciones importantes como la división celular, patogénesis, respuesta al estrés, inanición y resistencia a antibióticos. Tras clonar WhiB1 y suministrarlo en *trans*, se observó una reducción de la actividad de biodesulfuración en *Gordonia*, y en *R. erythropolis*, por lo que parece que actúa como represor. Esta idea se refuerza por el hecho de que la sobreexpresión de WhiB1 disminuye la transcripción de los genes *dsz* y de la proteína de la familia TetR descrita como regulador transcripcional del operón *dsz*. Sin embargo, se desconoce como se ejerce la represión; si es por unión directa al promotor, o mediada por el regulador transcripcional, o por ambas vías. (Murarka *et al.*, 2020).

Hasta el momento, no se ha descrito ningún tipo de regulación de los genes *dsz* que responda a DBT, por lo que es probable que no haya una respuesta específica a DBT, sino que la regulación de la expresión génica de los genes *dsz*, se englobe dentro de la respuesta general de la bacteria al estrés y a la limitación de azufre.

II.2. REGULACIÓN POR AZUFRE

Aunque el azufre sólo supone el 1% del total de la biomasa celular, es uno de los elementos esenciales para los organismos vivos, necesario para la síntesis de proteínas

y cofactores (Bordoloi *et al.*, 2016). La vía metabólica clave de asimilación de azufre por parte de los microorganismos es la biosíntesis de cisteína a partir del sulfato inorgánico que es la fuente preferencial de azufre para la mayoría de los microorganismos. Los genes que participan en la incorporación de azufre vía síntesis de cisteína forman parte del regulón *cys* y han sido caracterizados en *Escherichia coli* y *Salmonella entérica*.

En condiciones de limitación de azufre muchas bacterias expresan proteínas denominadas “sulfur starvation induced” (SSI) o “sulfur starvation response” (SSR) capaces de metabolizar fuentes alternativas de azufre. Estas proteínas se sintetizan exclusivamente en ausencia de fuentes preferenciales de azufre e incluyen sistemas de transporte y enzimas para liberar el azufre de los compuestos orgánicos. Las proteínas Dsz, podrían formar parte de esta categoría (Mohebbi & Ball, 2008).

El regulador principal de los genes de respuesta a limitación de azufre es CysB, existiendo además otros reguladores adicionales como Cbl en enterobacterias y SfnR en *Pseudomonas*. Por otra parte, la expresión de los genes activados por SfnR, requiere además un factor sigma alternativo.

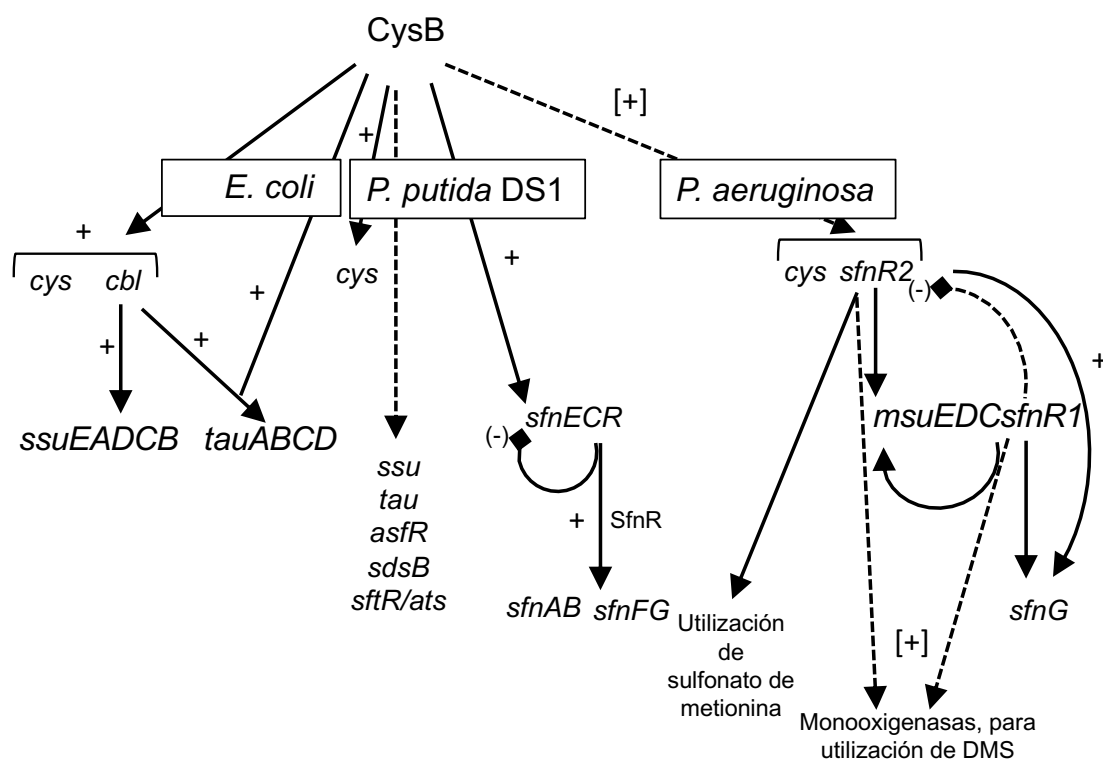


Figura 2: Representación esquemática de la regulación transcripcional de los genes SSI, en *E. coli*, *P. putida* DS1 y *P. aeruginosa*. El símbolo $+$ indica regulación positiva, el símbolo $-$ indica regulación negativa y las líneas punteadas indican posible relación.

La regulación transcripcional de los genes SSI, en *E. coli* está muy estudiada, y se esquematiza en la figura 2. Los operones mejor caracterizados son los implicados en la utilización de sulfonatos alifáticos (*ssuEADCB*) y taurina (*tauABCD*). Estos genes codifican para los transportadores tipo ABC (TauABC en el caso de taurina y SsuABC en el caso de sulfonatos alifáticos) y sus respectivas oxidorreductasas (TauD y SSuDE), que liberan sulfito. Puesto que el sulfito entra en la ruta biosintética clásica de la cisteína, la asimilación de los sulfonatos como fuentes de azufre depende en última instancia del regulador CysB, y puede considerarse como una extensión del regulón *cys*. CysB, es un regulador tipo LysR, activador transcripcional considerado como el “regulador maestro” de la regulación por fuente de azufre en bacterias, que está muy conservado en Gram negativas.

CysB, activa la transcripción de los genes *cys*, implicados en la asimilación de sulfato, (siendo en este caso el único activador implicado), y a *cbl*, que codifica para un segundo regulador transcripcional tipo LysR implicado en la expresión de los operones *ssuEADCB* y *tauABCD*. En *Salmonella enterica* serovar *Typhimurium*, CysB se une a zonas ricas en AT de los promotores de *cysJIH*, *cysK*, y *cysP*. Esta unión, que se produce en presencia de N-acetilserina o de O-acetilserina, está muy bien caracterizada; pero, sin embargo, no se ha podido definir una secuencia consenso para los sitios de unión de CysB (Kouzuma *et al.*, 2008). Los genes *cys* sólo están parcialmente reprimidos por sulfato, mientras que los sistemas *tau* y *ssu* están totalmente reprimidos tanto por cisteína como por sulfato (Bykowski *et al.*, 2002). Aunque la expresión de los genes *tau* y *ssu* está regulada por Cbl, la regulación de estos operones es diferente, ya que la expresión de los genes *tau* requiere la participación directa de CysB y Cbl sobre su región promotora reguladora, mientras que la expresión de los genes *ssu* requieren de CysB solo para que se transcriba *cbl*. (van der Ploeg *et al.*, 1999). La regulación transcripcional de los genes *tau* está muy bien caracterizada: van der Ploeg *et al.*, 1997, demostraron mediante análisis *in vitro* que la presencia de Cbl por si sola no es suficiente para activar el promotor *tau*, debido a la posición desfavorable de su sitio de unión con respecto a la región -10. Cbl unido a su sitio de unión ayudaría a reclutar a CysB, o reposicionarlo correctamente, permitiendo una eficiente transcripción. CysB y Cbl tienen una similitud de secuencia de aminoácidos excepcionalmente alta, incluidos sus motivos de unión de ADN de hélice-vuelta-hélice, lo que posiblemente permite este “cross talk”. (van der Ploeg *et al.*, 1997). La expresión de los genes regulados por CysB, está mediada por moléculas como la N-acetilserina o la O-acetilserina, que actúan como inductores internos señalizando la limitación de azufre, mientras que el sulfuro y el tiosulfato actúan como anti-inductores, compitiendo con la N-acetilserina e inhibiendo la unión de CysB a los promotores *cys*. En ausencia

de sulfato, Cbl se une a secuencias aguas arriba de los operones *ssu* y *tau*, activando su transcripción, sin necesidad de un ligando. En presencia de sulfato, la adenosina-5'-fosfosulfato (APS) (intermediario en la asimilación de sulfato), actúa como señal molecular co-represora de Cbl, inhibiendo la unión de Cbl al ADN y por tanto la expresión de los genes *ssu* y *tau*. (Kouzuma *et al*, 2008).

Aunque en *Pseudomonas* también existe CysB como regulador principal que dirige la respuesta a limitación de azufre, la regulación de la expresión de los genes SSI es diferente, puesto que no existe un gen equivalente a *cbl*. CysB controla de forma directa la expresión de los genes *cys*, de forma indirecta la respuesta mediada por las proteínas tipo SfnR, y la expresión de una serie de genes implicados en la utilización de azufre orgánico, que de alguna forma dependen de CysB, pero son independientes de SfnR. Se ha demostrado que CysB de *P. putida* DS1, bacteria que puede utilizar sulfuro de dimetilo (DMS) como fuente de azufre, se une también a zonas ricas en AT de la región aguas arriba del operón *sfnECR*, activando su expresión, pero tampoco se ha podido definir una secuencia consenso de unión a ADN.

El operón *sfnECR* codifica para una reductasa dependiente de NADH (SfnE), para una monooxigenasa (SfnC) y para un activador transcripcional dependiente de σ^{54} que no tiene dominio regulador N terminal (SfnR) (Endoh *et al.*, 2003). SfnR activa la expresión del operón *sfnFG* necesario para la utilización de dimetilsulfona (DMSO₂). SfnR también activa la expresión del operón *sfnAB* implicado en la utilización de metanotiol como fuente de azufre (Habe *et al.*, 2007, Kouzuma *et al*, 2008). Mediante la construcción de un mutante carente del factor σ^{54} , se puso de manifiesto que SfnR controla la expresión de los genes involucrados en el metabolismo de DMSO₂ en *P. putida* DS1 por interacción directa con dicho factor sigma alternativo (Endoh *et al*, 2003). Por otra parte, SfnR de *P. putida* DS1 se autorregula, uniéndose a la región aguas abajo del TSS de *sfnE* e impidiendo su transcripción. (Kouzuma *et al*, 2008).

La mayoría de los reguladores dependientes de σ^{54} tienen la misma arquitectura: un dominio de unión al ADN, (dominio C-terminal), un dominio central con actividad ATPasa que interacciona con el factor σ^{54} de la ARN polimerasa, y un dominio regulador (N-terminal). Sin embargo, como se ha comentado, SfnR no tiene dominio regulador. El activador transcripcional dependiente de σ^{54} sin dominio N terminal mejor caracterizado es la proteína PspF de *E coli*, que controla la expresión del operón *pspABCDE*, necesario para sobrevivir en fase estacionaria en ambiente alcalino. PspF tiene actividad constitutiva, activando el operón *psp* sin necesidad de fosforilación ni de unión a ligando, sino que su control está mediado por la formación de un complejo represor

con la proteína PspA. Es posible que SfnR esté regulada negativamente de manera similar, por unión a una proteína asociada en presencia de sulfato (Kouzuma *et al*, 2008) pero hasta el momento no se ha descrito ninguna proteína reguladora que pudiera participar en la regulación junto con SfnR.

P. aeruginosa PAO1, *Acinetobacter baumannii* y *Enterobacter cloacae* tienen 2 genes que codifican para proteínas SfnR: SfnR1 y SfnR2. En *P. aeruginosa* PAO1 estas 2 proteínas poseen una similitud entre sí del 65%. SfnR1 está codificada en el operón *msuEDCsfnR1*, mientras que el gen *sfnR2* es monocistrónico. *P. aeruginosa* PAO1 puede utilizar DMS, DMSO y DMSO₂ como fuentes de azufre. En ausencia de fuentes preferenciales de azufre, o en limitación de azufre (pero no necesariamente en presencia de DMS o sus derivados) se induce la expresión del gen *sfnR2*, siendo dicha inducción muy probablemente mediada por CysB. Una vez producida, SfnR2 activa la transcripción de *msuEDCsfnR1* y de *sfnG*, desde promotores dependientes de σ^{54} (por lo que σ^{54} debe tener un papel en la regulación del sistema), situados aguas arriba de dichos genes. SfnR1 participa en la activación de *sfnG* y autorregula la suya propia, a la vez que regula negativamente a *sfnR2*. Además, *sfnR2* también es necesario para la utilización de metionina sulfona como fuente de azufre. A pesar de que SfnR1 es muy similar a SfnR de *P. putida* (80% de identidad), y de que SfnR2 sólo presenta un 60% de identidad con otros SfnR, SfnR2 es el regulador principal implicado en la utilización de DMS como fuente de azufre, siendo esencial para este proceso. Además, es muy probable que SfnR2 y SfnR1 regulen otros genes implicados en la utilización de DMS, incluyendo las monooxigenasas necesarias en la conversión de DMS a DMSO₂. (Lundgren *et al.*, 2018)

La activación de los genes inducidos en respuesta a limitación de azufre que son independientes de SfnR en las distintas especies del género *Pseudomonas*, puede estar mediada por 3 activadores tipo LysR diferentes codificados por los genes *sdsB*, involucrado en la utilización de SDS en *Pseudomonas* sp. ATCC19151 (Davison *et al*, 1992), *asfR*, que modula la transcripción del operón *asfABC*, necesario para la utilización del sulfonato de arilo en *P. putida* S-313 (Vermeij *et al.*, 1999) y *sftR*, que activa la expresión del grupo de genes *atsBC*, *atsRK* y *sftP-astA*, cuyos productos génicos están involucrados en la utilización de sulfonatos de arilo o de alquilo en *P. putida* S-313 (Kahnert *et al.*, 2002). Por otra parte, la regulación transcripcional de los genes *ssu* y *tau*, tampoco parece estar controlada por SfnR en *Pseudomonas*, aunque se desconoce cómo se regulan (Endoh *et al*, 2003).

Los genes *ssu* y los genes *seu* de *Corynebacterium glutamicum* ATCC13032 se expresan mucho más cuando el crecimiento se produce en sulfonatos y ésteres de sulfonato, que cuando se produce en sulfato. Esta regulación también parece responder

a la disponibilidad de azufre y no a una fuente de azufre concreta, ya que *ssuD1* parece estar regulada por el represor McbR, que es un regulador global del metabolismo del azufre en esta bacteria. (Koch *et al.*, 2005)

Un ejemplo interesante de genes regulados en respuesta a limitación de azufre, son los implicados en la producción de biosurfactantes, los cuales se expresan más en condiciones de limitación de azufre. En presencia de DBT como única fuente de azufre las células sufren estrés por la baja solubilidad del DBT y producen más biosurfactantes, probablemente para facilitar la asimilación de fuentes de azufre de menor solubilidad, como DBT, lo cual es muy interesante, porque se puede emplear para mejorar el proceso de BDS. Con ensayos de RT-qPCR se ha comprobado que en *Pseudomonas* sp. AK6U la transcripción de los genes *rhlABC*, que codifican para la enzima clave en la biosíntesis de dirhamnolípidos, depende de la fuente de azufre, ya que *rhlABC* está regulado positivamente en DBT y DBTO₂ indicando que la bacteria produce más dirhamnolípidos en estas fuentes de azufre que en MgSO₄ (Ismail *et al.*, 2014).

Por otra parte, puesto que σ^{54} tiene un papel en el metabolismo del azufre, participando en el control de la expresión de los genes *sfn* en *P. putida*, y además también participa en la regulación de la producción de biosurfactantes en *P. aeruginosa*, es posible que los factores sigma alternativos sean la conexión entre el metabolismo del azufre y la producción de biosurfactantes (Ismail *et al.*, 2014).

Todos estos antecedentes sugieren, por tanto, que la expresión de los genes de degradación de DBT forme parte de la respuesta general a la disponibilidad de azufre en vez de una inducción por la fuente de azufre concreta.

III. METAGENÓMICA FUNCIONAL

Aunque en las 2 últimas décadas se han descrito gran cantidad de genes similares al operón *dszABC* de *Rhodococcus erythropolis* IGTS8 en diferentes bacterias, siempre se ha tratado de estudios basados en organismos cultivables, los cuales representan un pequeño porcentaje de la diversidad bacteriana, limitando mucho el potencial biodesulfurador susceptible de ser estudiado (Vilchez-Vargas *et al.*, 2010). Se estima que tan sólo el 0,1-1 % de los microorganismos pueden ser cultivados con las técnicas actuales, por lo que nuevas técnicas, entre ellas los métodos metagenómicos, pueden ser utilizados para explorar nuevos genes y vías metabólicas para la biodegradación de moléculas recalcitrantes y xenobióticas (Xu *et al.*, 2006).

La mayoría de las enzimas que se emplean actualmente en la industria como catalizadores biológicos son de origen microbiano. Esto es debido a la enorme diversidad metabólica de los microorganismos, que constituyen el mayor reservorio de diversidad genética de la biosfera. Para acceder a este potencial genético, se han desarrollado enfoques de “metagenómica ambiental”. La metagenómica consiste en la construcción y selección de bibliotecas de genomas que no proceden de un único microorganismo, sino de toda la comunidad microbiana de muestra ambiental. (Ferrer *et al.*, 2009). Para la construcción de la biblioteca, es importante tener en cuenta cuál es la actividad enzimática de interés, para elegir zonas de muestreo de lugares potencialmente enriquecidos en dicha actividad, ya que la construcción de bibliotecas metagenómicas a partir de estos sitios puede aumentar las probabilidades de obtener resultados positivos. Por ejemplo, para encontrar enzimas de degradación de contaminantes xenobióticos, los sitios de muestreo deben estar contaminados con el compuesto de interés o compuestos relacionados (Terrón-González *et al.*, 2014). De hecho, se han llevado a cabo con éxito estudios de metagenómica funcional con el objetivo de encontrar enzimas con actividades degradadoras de contaminantes aromáticos, como actividad extradiol dioxigenasas a partir de muestras de lugares contaminados, (Brennerova *et al.*, 2009, Terrón-González *et al.*, 2016).

Hay 2 grandes grupos de estudios metagenómicos: los basados en el análisis de secuencias, y los basados en el análisis de función, o metagenómica funcional. Los análisis basados en secuencia consisten en secuenciación aleatoria de ADN metagenómico, que puede hacerse a partir de clones de una metagenoteca o directamente del ADN metagenómico purificado, y que ha evolucionado mucho en los últimos años gracias al desarrollo de las técnicas de secuenciación masiva de nueva generación (NGS, *Next Generation Sequencing*), la cual ha permitido la realización de gran número de estudios metagenómicos. Sin embargo, tiene la limitación de que a menudo se generan lecturas cortas, lo que hace más difícil el ensamblaje en secuencias más largas (contigs). Por otra parte, al generarse una gran cantidad de secuencias se dificulta la realización de correcciones manuales, y las anotaciones quedan limitadas a métodos automáticos. (Terrón-González, 2014). De hecho, aunque la secuenciación de ADN ambiental es una aproximación interesante para obtener información de microorganismos no cultivados, debido a la alta complejidad de las comunidades microbianas y el pequeño tamaño de la mayoría de las lecturas, pueden producirse ensamblajes artificiales de secuencias de ADN derivadas de diferentes hospedadores (Ferrer *et al.*, 2009).

La ventaja principal de los estudios metagenómicos basados en secuencia es que no se necesita la expresión de genes clonados en hospedadores que pueden diferir mucho de su fondo genético original. Mientras que su mayor inconveniente es que el descubrimiento de nuevas enzimas está limitado a aquellas que tengan similitud con otras conocidas, de forma que se pueda establecer una relación entre la nueva secuencia y alguna contenida en las bases de datos.

Un ejemplo de este tipo de estudios es el llevado a cabo por Abbasian *et al.*, (Abbasian *et al.*, 2016) en el que se describen nuevos genes *dsz* de bacterias no cultivadas. A partir de un estudio metagenómico de suelos de campos de petróleo, se identificaron un gran número de fragmentos de genes responsables de la desulfuración de DBT y aunque la mayoría eran similares a genes ya existentes en las bases de datos, también se descubrieron secuencias nuevas, a partir de las cuales se consiguieron obtener las secuencias completas de los genes, con técnicas bioinformáticas como el empleo del software “Geneious” y el mapeo reiterativo de lecturas cortas de todos los datos metagenómicos (Abbasian *et al.*, 2016).

En cuanto a los análisis basados en función, permiten recuperar clones positivos de una biblioteca metagenómica sin información de secuencia previa, con lo que se puede identificar genes codificantes de enzimas con actividades conocidas pero completamente distintos a los existentes en las bases de datos; pero además, permite vincular fenotipos promiscuos a secuencias (Fernández-Arrojo *et al.*, 2010), aumentado por tanto el espectro de actividades y reacciones enzimáticas, susceptibles de ser encontradas y posteriormente utilizadas. Un ejemplo de promiscuidad catalítica es la existencia de enzimas con diferentes mecanismos catalíticos en el mismo sitio activo. Esta promiscuidad puede ser resultado de la evolución natural, ya que mejora la flexibilidad metabólica del microorganismo y su adaptabilidad al ambiente; o resultado de la evolución en el laboratorio, y, en cualquier caso, puede ser útil para numerosas aplicaciones. Beloqui *et al.*, 2010, demostraron la utilidad de la metagenómica para acceder a nuevas actividades catalíticas, ya que describieron, una misma enzima con actividad esterasa y deshalogenasa, que fue aislada de una metagenoteca construida a partir de agua de mar contaminada con petróleo crudo, (Beloqui *et al.*, 2010).

La metagenómica funcional, se basa en la detección de un cambio fenotípico en el hospedador, debido a la adquisición de un clon de la metagenoteca que posea una función o actividad particular. (Terrón *et al.*, 2014). Por lo que su principal problema es conseguir una expresión heteróloga suficiente para permitir la detección de dicha actividad. Para tratar de solventar este problema, hay que tener en cuenta tanto las características del vector como las del hospedador.

En cuanto al hospedador, generalmente se emplean diferentes estirpes de *E. coli* porque permite la replicación de todos los vectores utilizados para construir metagenotecas, y porque la mayoría de las características de interés de dichos vectores, como la alta eficiencia de empaquetamiento en partículas del fago lambda para la transfección posterior, solo son aplicables en *E. coli*. Por otra parte, el ADN puede ser fácilmente recuperado a partir de dicha bacteria. (Terrón *et al.*, 2014).

En el caso del vector, hay que tener en cuenta varios aspectos importantes como su sistema de replicación, nº de copias, o tamaño del inserto que puede alojar, pudiéndose emplear plásmidos convencionales y vectores derivados del fago lambda, para insertos de hasta 12-20 kb, BACs, o cromosomas artificiales bacterianos, para insertos de hasta 300 kb y los más empleados que son los cósmidos y particularmente los fósidos.

Los cósmidos son plásmidos que poseen sitios *cos*, por lo que pueden ser empaquetados en cabezas del fago lambda, lo cual mejora la frecuencia de incorporación al hospedador, ya que se realiza por transfección en lugar de por transformación. Es importante el tamaño del propio cósmido en sí, puesto que la suma del vector y del inserto no puede superar la capacidad de empaquetamiento en una cabeza del fago lambda (50-52 kb). Hay cósmidos, de amplio espectro de replicación, y otros de espectro reducido. Muchos son de alto número de copias, lo que causa inestabilidad a las bibliotecas metagenómicas sobre todo en el caso de insertos grandes.

Los fósidos son muy interesantes para la construcción de metagenotecas, porque reúnen ventajas de los cósmidos, ya que poseen el sitio *cos*, con lo que pueden ser empaquetados en cabezas de fago lambda, y de los BACs, puesto que poseen el sistema de replicación del plásmido F, con lo que se mantienen de forma estable en un bajo número de copias (1-2), aunque solo se puedan replicar en *E. coli*. (Terrón, 2014). En el caso de los análisis funcionales, hay que tener en cuenta que, si se van a aprovechar los promotores heterólogos que suele haber adyacentes al sitio de clonación en los vectores para forzar la expresión de genes metagenómicos, su capacidad para expresar de forma eficiente los genes metagenómicos clonados correlaciona inversamente con el tamaño del inserto, siendo un factor limitante importante la presencia de terminadores de la transcripción aguas arriba del gen de interés (Terrón *et al.*, 2013).

Por otra parte, también hay que tener en cuenta el número de copia de los vectores, ya que como se ha comentado con vectores de alto número de copias la estabilidad de la metagenoteca normalmente será menor. En el caso de los estudios funcionales, también hay que considerar que al aumentar el número de copias aumentará la dosis génica, lo que puede permitir detectar ciertas actividades cuya

expresión sea demasiado baja en la bacteria hospedadora, aunque también podría causar un efecto deletéreo, en el caso de la expresión de genes cuyos productos sean tóxicos para el hospedador, por lo que los vectores que permiten mantener la metagenoteca en baja copia e incrementar el número de copias durante el rastreo funcional son muy interesantes (Terrón, 2014).

Como se ha comentado, el principal problema es conseguir una expresión heteróloga suficiente para permitir la detección de dicha actividad. Aunque la expresión génica es un proceso de varios pasos que implica transcripción, traducción y modificaciones postraduccionales, la transcripción es el proceso regulado con más frecuencia y es probablemente el principal paso limitante de la expresión del ADN metagenómico (Terrón *et al.*, 2013). Terrón *et al.*, 2013 desarrollaron un vector para la construcción de metagenotecas así como 2 estirpes de *E. coli*, para la realización de rastreos funcionales, con los que se consigue aumentar la eficiencia de expresión de los genes metagenómicos, mejorando su transcripción. En cuanto al vector, a partir del fósido comercial pCC1FOS (Epicentre/Lucigen) que permite la inducción del número de copias con la adición de arabinosa, se construyó el vector pMPO579 que incorpora tres elementos importantes: el origen de transferencia conjugativa *oriT*, para permitir la transferencia de la biblioteca metagenómica por conjugación, un sistema de transcripción sujeto a anti-terminación (promotor P_{sal} seguido del sitio *nut* de reconocimiento de la proteína antiterminadora N del fago lambda) para mejorar la expresión de los genes metagenómicos, y un gen *gfp* sin promotor para permitir la detección de la expresión génica metagenómica. Este vector permite la transcripción del ADN metagenómico clonado por dos ARN polimerasas distintas: (i) la ARN polimerasa de T7, desde el promotor del *gen 10* de T7, ya presente en el vector original pCC1FOS (Epicentre/Lucigen), y (ii) la ARN polimerasa bacteriana modificada para la anti-terminación procesiva por la proteína N, desde el promotor P_{sal} . Para aprovechar las ventajas de estas modificaciones, se construyeron dos estirpes, MPO553 y MPO554, derivadas de la estirpe comercial EPI300-T1 (Epicentre/Lucigen). MPO553, cuyo sistema de expresión se basa en el de T7, posee el promotor lacUV5, el atenuador *nasF*, y el gen-1 del fago T7 inserto en su cromosoma, y produce niveles constitutivos pero bajos de la ARN polimerasa de T7. MPO554, que se ha empleado en este trabajo de Tesis, contiene el gen activador *nahR* y el promotor P_{sal} seguido por el gen *N* del fago lambda. Tras la inducción con salicilato, NahR activaría el P_{sal} del fósido, activando así la transcripción del ADN metagenómico, y simultáneamente el P_{sal} de la estirpe especializada empleada como hospedador, activando así la anti-terminación procesiva

dependiente de la proteína N en las transcripciones iniciadas desde P_{sal} del vector (Terrón *et al.*, 2013).

Con estas herramientas, la Dra. Terrón llevó a cabo un estudio de metagenómica funcional, realizando una búsqueda de clones con actividad oxigenasa hidroxilante de anillos aromáticos, u oxigenasa inicial, en una metagenoteca construida con muestras de suelo contaminado, procedente de una industria petroquímica. Para este tipo de rastreo se empleó como sustrato inicial el triptófano presente en el medio LB, a partir del cual, la triptofanasa endógena de *E. coli* es capaz de producir indol. El anillo heterocíclico del indol es susceptible de ser hidroxilado por muchas oxigenasas hidroxilantes obteniendo así índigo, de color azul. Con lo cual, los clones de la metagenoteca que tuvieran una actividad oxigenasa inicial capaz de utilizar indol como sustrato podrían ser detectados al colorearse de azul. Con este rastreo, se rescataron de la metagenoteca 25 clones diferentes, uno de los cuales (UPO21) contenía un nuevo operón degradador de DBT (Terrón, 2014).

UPO21 contiene, acoplado a los genes *dszABC*, el gen que codifica para la FMN reductasa de la ruta siendo la primera vez que se describe un gen de un transportador de electrones acoplado a un operón *dsz*. Además, el producto génico presenta más similitud con las FMN reductasas implicadas en la desulfonación de sulfonatos alifáticos (SsuE) que con las reductasas caracterizadas de la ruta de degradación de DBT (DszD). Teniendo en cuenta que este gen y *dszA* estaban separados únicamente por 4 nucleótidos, lo que sugiere acoplamiento traduccional y su implicación en la degradación de DBT, pero que tiene poca similitud con la reductasa DszD, se decidió nombrar su producto génico como DszE (Terrón, 2014) (Martín-Cabello *et al.*, 2020). En UPO21 se identificó además el posible activador transcripcional de la ruta, un activador dependiente de σ^{54} denominado DszR (Terrón, 2014), que se caracterizará con detalle en este trabajo de Tesis.

Estos resultados ponen claramente de manifiesto los beneficios de llevar a cabo estudios de metagenómica funcional para incrementar nuestro conocimiento de los procesos metabólicos y funcionales existentes en la naturaleza.

OBJETIVOS

OBJETIVOS

- I. Caracterización de los genes catabólicos *dszEABC***
- II. Caracterización de la regulación del operón.**
- III. Construcción de casetes de biodesulfuración de DBT de amplio espectro.**

MATERIALES Y MÉTODOS

I. Estirpes, plásmidos y oligonucleótidos.

Se relacionan a continuación, las estirpes, los plásmidos y los oligonucleótidos utilizados en la realización de este trabajo.

Tabla 1: estirpes empleadas en este trabajo.

| Estirpe | Características relevantes | Referencia |
|--|---|-----------------------------------|
| <i>E. coli</i> DH5 α | α : F'80d <i>lacZ</i> Δ M15 Δ (<i>lacZYA-argF</i>)U169 <i>recA1 endA1 hsdR17</i> ($r_k^- m_k^+$) <i>supE44 thi-1 gyrA relA1</i> | Hanahan, 1983 |
| <i>E. coli</i> EPI300-T1 ^R | F' <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) (Str ^R) Φ 80d <i>lacZ</i> Δ M15 Δ <i>lacX74 recA1 endA1 araD139</i> Δ (<i>ara, leu</i>)7697 <i>galU galK</i> $\lambda^- rpsL nupG trfA tonA dhfr$. | Epicentre. |
| <i>E. coli</i> MPO554 Nal ^R | Derivada de EPI300-T1 ^R , resistente natural a nalidíxico. | Terrón-González, 2014. |
| <i>E. coli</i> NCM631 | <i>hsdS gal lacZ</i> Δ M15: <i>lacI lacUV5::gene1</i> (ARNP T7) Δ <i>lac-Tn10</i> . | Govantes & Santero, 1996. |
| <i>Pseudomonas putida</i> DOT-T1E | Rif ^R , Tol ^R , Tol ⁺ | Ramos <i>et al.</i> , 1995 |
| <i>Pseudomonas putida</i> KT2440 | mt-2 <i>hsdR1</i> ($r^- m^+$). | Bagdasarian <i>et al.</i> , 1981. |
| <i>Pseudomonas putida</i> MPO406 | mt-2 <i>hsdR1</i> ($r^- m^+$) <i>cbrB::</i> MiniTn5 <i>luxABKm</i> ; Rif ^R , Km ^R | Amador <i>et al.</i> , 2010 |

Tabla 2: plásmidos empleados en este trabajo.

| Plásmido | Características relevantes | Referencia |
|----------|--|---------------------------------------|
| pBSKII+ | Vector de clonación. Ap ^R . | Stratagene |
| pIZ227 | Derivado LysE (Kleiner y col., 1998; Studier y Moffat, 1986), lleva el represor <i>lacI^q</i> , del plásmido pMM40. Cm ^R . | Govantes & Santero, 1996. |
| pIZ1016 | Vector de amplio espectro derivado del pBBRMCS-5, posee el promotor tac y el <i>lacI^q</i> del plásmido pMM40. Gm ^R . | Martínez-Pérez <i>et al.</i> , 2004. |
| pMPO200 | Vector de amplio espectro, para construir fusiones traduccionales al gen <i>lacZ</i> . Basado en pBBR1MCS-4. Ap ^R . | García-González <i>et al.</i> , 2005 |
| pMPO234 | Vector de amplio espectro, para construir fusiones transcripcionales al gen <i>lacZ</i> , basado en pMPO200. Ap ^R . | Porrúa <i>et al.</i> , 2009 |
| pMPO579 | Fósmido para construcción de metagenotecas, contiene <i>oriT</i> , promotor T7 y promotor <i>P_{sal}</i> , acoplado al sitio del fago lambda <i>nutL</i> . Cm ^R . | Terrón-González <i>et al.</i> , 2013. |
| pMPO587 | Derivado de UPO21, contiene los genes <i>dszE</i> , <i>dszA</i> , <i>dszB</i> y <i>dszC</i> , clonados en pMPO579. Cm ^R . | Terrón-González, 2014. |
| pMPO588 | Derivado de UPO21, contiene el gen <i>dszR</i> clonado en pIZ1016. Gm ^R . | Terrón-González, 2014. |
| pMPO593 | Fusión traduccional <i>dszEABC_{UPO21::lacZ}</i> , construida en pMPO579. Cm ^R . | Terrón-González, 2014. |
| pMPO597 | Fusión traduccional <i>dszEA_{UPO21::lacZ}</i> , construida en pMPO579. Cm ^R . | Terrón-González, 2014. |
| pMPO1388 | Fusión transcripcional <i>dszEABC_{UPO21::lacZ}</i> , construida en pMPO579. Cm ^R . | Este trabajo. |
| pMPO1390 | Fusión transcripcional <i>dszEA_{UPO21::lacZ}</i> , construida en pMPO579. Cm ^R . | Este trabajo. |

| | | |
|----------|--|---------------------------------|
| pMPO1394 | Fusión traduccional <i>dszE_{UPO21}::lacZ</i> , construida en pMPO579. Cm ^R | Este trabajo. |
| pMPO1395 | Fusión transcripcional <i>dszE_{UPO21}::lacZ</i> , construida en pMPO579. Cm ^R . | Este trabajo. |
| pMPO1396 | ΔP_{dsz} en pMPO1394, el +1 queda a la misma distancia de <i>P_{sal}</i> , de la que está originalmente de su propio promotor. Cm ^R . | Este trabajo. |
| pMPO1541 | ΔP_{dsz} en pMPO597. Cm ^R . | Este trabajo. |
| pMPO1544 | derivado de pTYB21, para superproducir y purificar DszR de UPO21 con una inteína fusionada en su extremo amino. Ap ^R . | Este trabajo. |
| pMPO1547 | Plásmido para obtener sondas para experimentos de retardo en gel y de protección frente a DNaseI. Fragmento de 167 pb que contiene los posibles sitios de unión de DszR y <i>P_{dsz}</i> (de UPO21) en pBluescript II SK+. Ap ^R . | Este trabajo. |
| pMPO1548 | Derivado de pTE103. Contiene el promotor <i>dsz</i> (de UPO21) e incluye el ATG (fragmento: de -160 a +42). Cuando se emplea como molde para transcripción in vitro se genera un transcrito de 342 pb (300 del plásmido). Ap ^R . | Este trabajo. |
| pMPO1550 | Fusión transcripcional <i>dszE_{UPO112}::lacZ</i> construida en el nucleótido 262 de <i>dszE</i> , en pMPO234. Ap ^R | Marica Noschese (no publicado). |
| pMPO1552 | Derivado de UPO112. Contiene <i>dszH</i> y <i>dszR</i> de UPO112, y 224 pb de secuencia aguas arriba de <i>dszH</i> , en pMPO579. Cm ^R | Julia Monserrat (no publicado). |
| pMPO1553 | Derivado de UPO112. Contiene <i>dszH</i> y <i>dszR</i> de UPO112, y 224 pb de secuencia aguas arriba de <i>dszH</i> en pIZ1016. Gm ^R | Este trabajo. |
| pMPO1554 | Deleción interna (1) en fase en <i>dszH</i> , se eliminan 188 codones de la parte central <i>dszH</i> . Conserva <i>P_{dsz}</i> y <i>dszR</i> . pIZ1016. Gm ^R | Este trabajo. |
| pMPO1556 | Deleción interna (2) en fase en <i>dszH</i> , se eliminan 176 codones de la parte central <i>dszH</i> . Conserva <i>P_{dsz}</i> y <i>dszR</i> . pIZ1016. Gm ^R . | Este trabajo. |
| pMPO1559 | ΔP_{dsz} en pIZ1016. Conserva <i>dszH</i> y <i>dszR</i> . Gm ^R . | Este trabajo. |
| pMPO1560 | ΔP_{dsz} en pMPO579. Conserva <i>dszH</i> y <i>dszR</i> . Cm ^R . | Este trabajo. |
| pMPO1561 | Deleción interna (1) en fase en <i>dszH</i> , se eliminan 188 codones de la parte central <i>dszH</i> . Conserva <i>P_{dsz}</i> y <i>dszR</i> . pMPO579. Cm ^R | Este trabajo. |
| pMPO1563 | ΔP_{dsz} y deleción interna (2) en fase en <i>dszH</i> , se eliminan 176 codones de la parte central <i>dszH</i> . Conserva <i>dszR</i> . pIZ1016. Gm ^R | Este trabajo. |
| pMPO1564 | ΔP_{dsz} y deleción interna (2) en fase en <i>dszH</i> , se eliminan 176 codones de la parte central <i>dszH</i> . Conserva <i>dszR</i> . pMPO579. Cm ^R | Este trabajo. |
| pMPO1565 | Genes <i>dszEABC</i> de UPO112, clonados en pBSKII+. Ap ^R . Tamaño: 8899 pb | Este trabajo. |
| pMPO1566 | Casete catabólico de DBT. Genes <i>dszHR</i> y <i>dszEABC</i> (incluye <i>P_{dsz}</i>), clonados en pSEVA221. Km ^R . Tamaño: 11837 pb | Este trabajo. |
| pMPO1570 | Casete catabólico de DBT. Contiene los genes <i>dszEABC</i> de UPO112, bajo el control de <i>alkS-P_{alkB}</i> . pSEVA429. Str ^R . Tamaño: 11236 pb | Este trabajo. |
| pMPO1571 | ΔP_{dsz} y $\Delta dszR$. Conserva <i>dszH</i> y 224 pb de <i>dszR</i> (incluyendo el codón de inicio). pIZ1016. Gm ^R . | Este trabajo. |
| pRK2013 | Plásmido auxiliar para conjugaciones. Replicón ColE1. Tra ⁺ , Km ^R . | Figurski & Helsinki, 1979. |
| pRK2073 | Plásmido auxiliar para conjugaciones. Str ^R , Spc ^R | Leong <i>et al.</i> , 1982. |

| | | |
|----------|--|--|
| pSEVA221 | Vector de amplio espectro, lleva los genes <i>oriV</i> , <i>trfA</i> y <i>oriT</i> . Km ^R . | Silva-Rocha <i>et al.</i> , 2013. |
| pSEVA429 | Vector de amplio espectro, lleva los genes <i>oriV</i> , <i>trfA</i> y <i>oriT</i> , y el sistema de expresión <i>alkS_PalkB</i> . Str ^R . | Silva-Rocha <i>et al.</i> , 2013. |
| pTE103 | Vector para transcripción <i>in vitro</i> , derivado del plásmido pUC18, sin diana HindIII. Ap ^R . | Elliott & Geiduschek, 1984. |
| pTYB21 | Vector para construir fusiones en el extremo amino terminal Ap ^R . | IMPACT TM Protein Purification System, NEB. |
| UPO21 | Fósmido con ADN metagenómico. Contiene los genes de desulfuración de DBT. Aislado en rastreo de clones con actividad productora de índigo. Cm ^R . | Terrón-González, 2014. |
| UPO112 | Fósmido con ADN metagenómico. Contiene los genes de desulfuración de DBT. Aislado en rastreo de clones con capacidad de crecer en DBT como única fuente de azufre. Cm ^R . | Este trabajo |

Se indican a continuación los detalles de la construcción de los diferentes plásmidos.

pMPO1388: Fusión transcripcional *dszEABC_{UPO21}::lacZ*, construida a partir de pMPO587, en el nucleótido 1241 de *dszC*. Se amplificó por PCR un fragmento de 650 pb con los oligonucleótidos EcoKpndszCfw y BamKpndszCrv; el cual se clonó en pMPO234 digerido con *EcoRI* y *BamHI*. Este plásmido intermedio se digirió con KpnI y Sall, y el inserto de 4,1 kb se hizo romo con T4 y se ligó con pMPO587 digerido con SphI. Tamaño: 18761 pb. Cm^R.

pMPO1390: Fusión transcripcional *dszEA_{UPO21}::lacZ*, construida a partir de pMPO587, en el nucleótido 464 de *dszA*. pMPO234 se digirió con *EcoRI* y Sall y el inserto de 3,5 kb se rellenó con klenow y se ligó con el fragmento de 11,4 kb obtenido digiriendo pMPO587 con XhoI y SphI y hecho romo con T4. Tamaño: 14908 pb. Cm^R.

pMPO1394: Fusión traduccional *dszE_{UPO21}::lacZ*, construida en el nucleótido 262 del gen. Se digirió pMPO587, con *BamHI* y Stul y el inserto de 1,13 kb se rellenó con Klenow. Este fragmento se ligó en pMPO200, cortado con XmaI y rellenó con Klenow. Este plásmido intermedio se cortó con *EcoRI* y Sall y el fragmento de 4,8 kb se rellenó con klenow y se ligó en pMPO579 cortado con SphI y hecho romo con T4. Tamaño: 14,1 kb. Cm^R.

pMPO1395: Fusión transcripcional *dszE_{UPO21}::lacZ*, construida en el nucleótido 262 del gen. El fragmento romo de 1,13 kb obtenido a partir de pMPO587 y empleado en la construcción de pMPO1394, se ligó en pMPO234 cortado con XmaI y rellenó con Klenow. El plásmido intermedio se cortó con *EcoRI* y Sall, y el fragmento obtenido (de 4,8 kb) se rellenó con klenow y se ligó en pMPO579 cortado con SphI y hecho romo con T4. Tamaño: 14,4 kb. Cm^R.

pMPO1396: Fusión traduccional *dszE_{UPO21}::lacZ* derivada de pMPO1394 con el promotor *dsz* delecionado. El +1 queda a la misma distancia de *P_{sal}*, de la que está originalmente de su propio promotor. Se deleciona un fragmento de 1209pb, que incluye además de *P_{dsz}*, el *P_{t7}* y el sitio *nut*. Se amplificó una banda de 1,77 kb a partir de pMPO579 con los oligonucleótidos: del pdszE-Fw y Rv; y se ligó con una banda de 11,4 kb obtenida digiriendo pMPO1394 con *Apal* y *Nsil*. Tamaño: 13 kb. Cm^R.

pMPO1541: Fusión traduccional *dszEA_{UPO21}::lacZ* (construida en el nucleótido 464 de *dszA*) derivada de pMPO597 con el promotor *dsz* delecionado. El +1 queda a la misma distancia de *P_{sal}*, de la que está originalmente de su propio promotor. Se deleciona un fragmento de 1209pb, que incluye además de *P_{dsz}*, el *P_{t7}* y el sitio *nut*. Se amplificó una banda de 1,77 kb a partir de pMPO579 con los oligonucleótidos: del pdszE-Fw y Rv; y se ligó con una banda de 11,4 kb obtenida digiriendo pMPO597 con *Apal* y *Nsil*. Tamaño: 13,7 kb. Cm^R.

pMPO1544: *dszR* de UPO21 clonado en pTYB21 (inteína fusionada en el extremo aminod de DszR). Se construyó por PCR, amplificando un framento de UPO21, de 1,1kb con los oligonucleótidos: PurifdszR Fw y Rv. El producto de PCR se digirió con *SapI* y *EcoRI* y se ligó con pTYB21 digerido con las mismas enzimas. El plásmido se secuenció completamente, para asegurar que no hubiera ninguna mutación. Tamaño: 8,6 kb. Ap^R.

pMPO1547: construido por PCR con los oligonucleótidos: *dszE* retFw y Rv, que amplifican 167 pb de pMPO1394 (contiene los posibles sitios de unión de DszR y *P_{dsz}*). El producto de PCR se digirió con *HindIII* y *Sall* y se ligó con pBSKII+ digerido con las mismas enzimas. El plásmido se secuenció completamente, para asegurar que no hubiera ninguna mutación. Tamaño: 3 kb. Ap^R.

pMPO1548: derivado de pTE103. Construido por PCR a partir de pMPO1394, con los oligonucleótidos *dszE* retFw y *dszE* IVTRv, que amplifican 212 pb que incluyen el promotor *dsz_{UPO21}* y el ATG de *dszE_{UPO21}* (fragmento: de -160 a +42). El producto de PCR se digirió con *HindIII* y se clonó en pTE103, digerido con la misma enzima. El plásmido resultante se secuenció para asegurar la orientación correcta y la ausencia de mutaciones. Tamaño: 3,2 kb. Ap^R.

pMPO1553: Derivado de UPO112. Contiene *dszH* y *dszR* de UPO112 clonado en plZ1016. Se construyó por restricción a partir de pMPO1552, digiriendo con *PmlI* y *HindIII*, el fragmento de 2047 pb, se rellenó con *klenow* y se clonó en plZ1016 digerido con *SmaI*. Se analizaron varios candidatos por restricción, para seleccionar el que tuviera *dszHR* bajo *P_{tac}*. Tamaño: 8,09 kb. Gm^R.

pMPO1554: Deleción interna (1) en fase en *dszH*, se eliminan 188 codones de la parte central *dszH*, conserva 45 pb incluyendo el ATG, y 42 pb incluyendo el codón de stop. Construido por PCR amplificando 2 fragmentos.

El inserto A (de 288 pb); se generó amplificando ADN de pMPO1552 con los oligonucleótidos: del dszS Fw-A y del dszS Rv-A y el producto de PCR se digirió con las enzimas *Sall* y *Nsil*.

El inserto B (de 1231 pb); se generó amplificando ADN de pMPO1552 con los oligonucleótidos: del dszS Fw-B y dszSR-112 Rv, y se cortó con las enzimas *Nsil* y *HindIII*.

Ambos insertos se ligaron entre sí y con pIZ1016 digerido con *Sall* y *HindIII*. Tamaño: 7523 pb. Gm^R.

pMPO1556: Deleción interna (2) en fase en *dszH*, se eliminan 176 codones de la parte central *dszH*, conserva 60 pb incluyendo el ATG, y 63 pb incluyendo el codón de stop. Construido por PCR amplificando 2 fragmentos.

El inserto A (de 303 pb); se generó amplificando ADN de pMPO1552 con los oligonucleótidos: del dszS Fw-A (el mismo empleado para construir pMPO1554) y del dszS Rv-A2, este producto de PCR se digirió con *Sall* y *Nsil*.

El inserto B (de 1246 pb); se generó amplificando ADN de pMPO1552 con los oligonucleótidos del dszS Fw-B2 y dszSR-112 Rv y se cortó con las enzimas *Nsil* y *HindIII*.

Ambos insertos se ligaron entre sí y con pIZ1016 digerido con *Sall* y *HindIII*. Tamaño: 7,544 kb. Gm^R.

pMPO1559: ΔP_{dsz} en pIZ1016. Construido por PCR amplificando un inserto *dszHR* a partir de pMPO1552, que lleva sólo 13 pb del gen anterior a *dszH* (inserto de 1,83 kb), con los oligonucleótidos del PdszS Sal Fw y dszSR-112 Rv, el cual se digirió con *Sall* y *HindIII* y se ligó con pIZ1016 digerido con *Sall* y *HindIII*. Tamaño: 7,871 kb. Gm^R.

pMPO1560: ΔP_{dsz} en pMPO579. Estrategia similar a la empleada en la construcción de pMPO1559. El inserto de 1,83 kb que contiene *dszH* y *dszR*, se generó con los oligonucleótidos del PdszS Pml Fw y dszSR-112 Rv. El producto de PCR se digirió con *PmlI* y *HindIII* y se ligó con pMPO579 digerido con *PmlI* y *HindIII*. Tamaño: 11363 pb. Cm^R.

pMPO1561: Se construyó siguiendo la misma estrategia que pMPO1554, pero el oligonucleótido Fw empleado para generar el inserto A, fue dszSR-112 Fw2, ya que la diana de restricción necesaria era diferente.

El inserto A se digirió con *PmlI* y *Nsil*, y se ligó con el inserto B, digerido con *Nsil* y *HindIII*, y con pMPO579, digerido con *PmlI* y *HindIII*. Tamaño: 11046 pb. Cm^R.

pMPO1563: ΔP_{dsz} y deleción interna (2) en fase en *dszH*, se eliminan 176 codones de la parte central *dszH*, conserva 60 pb incluyendo el ATG, y 63 pb incluyendo el codón de stop.

Se construyó por PCR, empleando como molde pMPO1556, y amplificando un fragmento de 1,83 kb con los oligonucleótidos delPdszS Sal Fw y dszSR-112 Rv, que se digirió con *Sall* y *HindIII* y se clonó en plZ1016 cortado con las mismas enzimas. Tamaño: 7349 pb. Gm^R.

pMPO1564: ΔP_{dsz} y delección interna (2) en fase en *dszH*, se eliminan 176 codones de la parte central *dszH*, conserva 60 pb incluyendo el ATG, y 63 pb incluyendo el codón de stop. Se construyó por restricción, digiriendo pMPO1563 con *HindIII* y *Sall*, el inserto de 1,323 kb se rellenó con *klenow* y se clonó en pMPO579 cortado con PmlI. Los candidatos obtenidos se evaluaron por restricción para seleccionar el que tuviera la construcción en la orientación correcta (bajo P_{sal}). Tamaño: 10942 pb. Cm^R.

pMPO1565: Es una construcción intermedia, para facilitar la posterior construcción de pMPO1566. Contiene los genes *dszEABC* de UPO112, clonados en pBSKII+. Se construyó por restricción, digiriendo UPO112 con *MluI*, *XhoI* y *NdeI*. Los fragmentos obtenidos se separaron mediante electroforesis, y se rescató el inserto de 5,9 kb, que posteriormente se rellenó con *klenow*. Este inserto se ligó con pBSKII+ digerido con *EcoRV* y se comprobó la orientación de los candidatos resultantes, analizándolos por restricción. Tamaño: 8899 pb. Ap^R.

pMPO1566: Contiene los genes reguladores (*dszHR*) y los genes catabólicos *dszEABC* de UPO112, clonados en pSEVA221. Se construyó en dos pasos.

1. Se digirió pMPO1565 con *HindIII* y *XbaI*, el inserto de 5,9 kb, se ligó con pSEVA221 digerido con las mismas enzimas. Este fragmento incluye 1082 pb del gen aguas arriba de *dszE*.

2. Los genes reguladores se obtuvieron por PCR a partir de UPO112, con los oligonucleótidos *dszR_orf1 HindFw* y *dszR_orf1Rv*, que amplifican 2085 pb. El inserto obtenido se digirió con *HindIII*.

El inserto del paso 2, se clonó en el plásmido construido en el paso 1, digerido previamente con *HindIII*. Se comprobó la orientación por restricción y se secuenció uno de los candidatos obtenidos (para descartar la presencia de mutaciones en la parte construida por PCR). Tamaño: 11837 pb. Km^R.

pMPO1570: Contiene los genes *dszEABC* de UPO112, bajo el control de *alkS_PalkB*. Se construyó por PCR, amplificando un fragmento de 4,36 kb empleando como molde pMPO1565, con los oligonucleótidos *Alk-dsz Fw* y *Alk-dsz Rv*. El producto de PCR se digirió con *HindIII* y *SpeI* y se clonó en pSEVA429, digerido con las mismas enzimas. Tamaño: 11236 pb. Str^R.

pMPO1571: Plásmido para complementar en *trans dszH*. ΔP_{dsz} y $\Delta dszR$ (conserva 224 pb de *dszR*, incluyendo el codón de inicio). Se construyó por restricción cortando

pMPO1559 con *EcoNI*, y con *HindIII*, el fragmento de 6,9 kb se rellenó con *klenow* y se religó. Tamaño: 6,992 kb. Gm^R.

Tabla 3: oligonucleótidos empleados en este trabajo.

| Nombre | Secuencia (5'- 3') |
|------------------|---|
| EcoKpndszCfw | TTAGAATTCGGTACCATGCGCCAGACCGACAG |
| BamKpndszCrv | TTAGGATCCCAGGAGGTGAATCCCGGCA |
| del pdszE-Fw | ATAACTGGCCTGAATATTC |
| del pdszE-Rv | GGTATGCATAACGATAACGGAGCAAAC |
| PurifdszR Fw | GGTGGTTGCTCTTCCAACATGAAATCGAAGATCGATCCG |
| PurifdszR Rv | GGTGGTGAATTCTCACCCCCGGCGCGACG |
| PEX-DBT1 | CGTTCGAGTTCTGTGCGGAAATGATCGAGC |
| SeqFpDszE1 | TCGACCAGATCCAATATGC |
| dszE retFW | GGTGGTAAGCTTATCGTGCCGGCGAGGCTTG |
| dszE retRV | GGTGGTGTGCGACCAGATCCAATATGCATAAAGCAG |
| dszE IVTRv | GAACAAGCTTTGCCATGACG |
| dszSR-Fw2 | GGTAAGCTTTCCGGCATGACGCATCGGC |
| dszSR-Rv2 | GGTTCTAGATGCAAGCGGTGCCTGCGTC |
| dszEUPO112Fw | GGTGGTAGGCCTGCTCAGGCGATCGCTAATG |
| dszEUPO112Rv | CGATAAAGGCCTTGAGCAAG |
| dszSR-112 Fw2 | ATATCACGTGCATGGCTGGACTCGTCTGG |
| dszSR-112 Rv | GGTGGTAAGCTTAGACCTGCAGACCTGAAACG |
| del dszS Fw-A | CGGGTCGACCATGGCTGGACTCGTCTGG |
| del dszS Rv-A | CGGATGCATTTTGGCCTGAAGCGCATTG |
| deldszS Fw-B | CGGATGCATCCATTCCCCAGCCAATTG |
| deldszS Rv-A2 | CGGATGCATGGGCACGATCAGAATTTTGG |
| deldszS Fw-B2 | CGGATGCATGACTCACCATCCGCCAATAC |
| delPdszS Sal Fw | CGGGTCGACGCAGGACCAATCCGATGCAA |
| delPdszS Pml Fw | CGGCACGTG GCAGGACCAATCCGATGCAA |
| dszR_orf1 HindFw | CGGAAGCTTAGACCTGCAGACCTGAAACG |
| dszR_orf1Rv | CGGAAGCTTTGGTTCGATGGCGAAACGCAC |
| Alk-dsz Fw | GGTAAGCTTGCTTTTTAGGCTGTGTCTTTGC |
| Alk-dsz Rv | GGTACTAGTGGAAGCTCGATACAGATCAGG |
| dszRQ Fw | AACACTCAGCCTCGACATTC |
| dszRQ Rv | CAAGCAGTTCCTGGGAACG |
| dszEQ Fw2 | CAACTGTCCTTGGATGGCTG |
| dszEQ Rv2 | ACGATCGCATCGGCTTCTTC |

II. Técnicas microbiológicas:

II.1. Medios y condiciones de cultivo.

Durante esta Tesis se ha trabajado con diferentes estirpes de *E. coli* y de *Pseudomonas*. Como **medio rico** para todas ellas se ha empleado el medio nutritivo Luria-Bertani (LB) (Sambrook & Russell, 2001).

LB: 10 g/l de triptona, 5 g/l de extracto de levadura y 5 g/l de cloruro de sodio. Agar bacteriológico americano 15 g/l, en el caso de medio sólido.

Como **medio mínimo** para las estirpes de *E. coli* EPI300-T1^R y derivadas, se ha utilizado una variante de M9 (Sambrook & Russell, 2001).

Variante de M9: Solución de sales 10 veces concentrada (Na_2HPO_4 423 mM, KH_2PO_4 220 mM, NH_4Cl 187 mM y NaCl 86 mM), pH 7,4. CaCl_2 0,1 mM. Glucosa 0,2 %. Leucina 0,5 mM. Isoleucina 0,3 mM. En los casos en que la fuente de azufre era diferente a MgSO_4 , se añadía además MgCl_2 0,1 mM.

Las diferentes fuentes de azufre empleadas, todas a 0,2 mM (salvo que se indique lo contrario) son: MgSO_4 , Taurina, DBT y 4,6 dimetil-DBT (estos 2 últimos disueltos en NN-DMF).

Para comprobar que la inducción de los genes de DBT no se debía a crecimiento pobre, se crecieron los plásmidos en medio con limitación de nitrógeno. Para ello, se preparó la misma solución de sales de la variante de M9, pero sin NH_4Cl . Las fuentes de nitrógeno en este caso fueron L-glutamina (1 mM), L-arginina (1 mM) y NH_4Cl (18,7 mM).

Agar noble 20 g/l, en el caso de medio sólido.

Se han utilizado 2 **medios mínimos** para las estirpes de *Pseudomonas*.

Medio mínimo para **experimentos de toxicidad:**

Se empleó un medio modificado a partir de Mandelbaum *et al.*, 1993, preparado a partir de las siguientes mezclas concentradas: tampón fosfato salino (solución 10 veces concentrada: 70 g/l de $\text{Na}_2\text{HPO}_4 \times 2 \text{H}_2\text{O}$, 28 g/l de KH_2PO_4 y 5 g/l de NaCl), una solución de microelementos (solución 500 veces concentrada: 50 g/l de $\text{MgSO}_4 \times 7\text{H}_2\text{O}$, 5g/l de $\text{FeSO}_4 \times 7\text{H}_2\text{O}$, 2,5 g/l de $\text{MnSO}_4 \times \text{H}_2\text{O}$, 3,2 g/l de ZnCl_2 , 0,33 g/l $\text{CaCl}_2 \times 2 \text{H}_2\text{O}$, 0,18 g/l de $\text{CuSO}_4 \times 5 \text{H}_2\text{O}$, 0,15 g/l de $\text{CoCl}_2 \times 6 \text{H}_2\text{O}$, 3,25 g/l de H_3BO_3 , 5 g/l de EDTA y 73 ml de HCl al 37%; filtrados y conservados a temperatura ambiente) y vitaminas (solución 100 veces concentrada: 10 mg/l de tiamina-HCl, 4 mg/l de ácido fólico, 20 mg/l de nicotinamida y 20 mg/l de piridoxina-HCl; esterilizado por filtración y conservado a 4 °C). Como fuente de carbono se utilizó succinato sódico 25 mM (preparado a 1M, autoclavado y conservado a temperatura ambiente). Se utilizó cloruro de amonio 1 g/l

(preparado a 100 g/l, autoclavado y conservado a temperatura ambiente) como fuente de nitrógeno. Agar noble 20 g/l, en el caso de medio sólido.

Medio mínimo para **experimentos de crecimiento en DBT como única fuente de azufre (medio M9/A9)**:

El medio mínimo anteriormente descrito, permite a *Pseudomonas* crecer a expensas del azufre acumulado, cuando se encuentra en limitación de azufre, lo que dificultaba los experimentos de crecimiento en DBT. Por eso se optó por un medio más pobre, modificado de Abril *et al.*, 1989. El cual llevaba además de la solución de sales 10 veces concentrada de la variante de M9; 2,5 ml/l de solución de microelementos A9 (0,3 g/l H_3BO_3 , 0,05 g/l ZnCl_2 , 0,03 g/l $\text{MnCl}_2 \times 4 \text{H}_2\text{O}$, 0,37 g/l $\text{CoCl}_2 \times 6 \text{H}_2\text{O}$, 0,010 g/l $\text{CuCl}_2 \times 2 \text{H}_2\text{O}$, 0,02 g/l $\text{NiCl}_2 \times 6 \text{H}_2\text{O}$ y 0,03 g/l $\text{NaMoO}_4 \times 2 \text{H}_2\text{O}$), CaCl_2 0,1 mM, como fuente de carbono se utilizó succinato sódico 25 mM y como fuente de azufre, MgSO_4 o DBT 0,2 mM. Agar noble 20 g/l, en el caso de medio sólido.

Los antibióticos utilizados, cuando fue necesario fueron: cloranfenicol (Cm) a 12,5 µg/ml para *E. Coli* y a 15 µg/ml para *Pseudomonas*, ácido nalidíxico (Nal) 15 µg/ml, kanamicina (Km) a 25 µg/ml, ampicilina (Ap) a 100 µg/ml, rifampicina (Rif) 20 µg/ml, carbenicilina (Cb) 100 µg/ml, tetraciclina (Tc) 5 µg/ml, Streptomycin (Str) a 50 µg/ml y gentamicina (Gm) 10 µg/ml.

Otros compuestos que se añadieron al medio cuando fue necesario son: arabinosa 1 mM, (para aumentar el nº de copias de los fósmodos), salicilato 0,5 mM-1mM (para inducir el promotor P_{sal}) IPTG 1mM (para inducir el promotor P_{tac}), X-gal 25 mg/l (para detectar actividad β-galactosidasa en medio sólido). 2-HBF (P28263, Sigma) preparado en NNDMF, DCPK 0,05% v/v (8.40157 merck) (para inducir el sistema $alkS_P_{alkB}$).

Los cultivos de *E. coli* se incubaron a 30 ó 37 °C, (según el experimento), y a 30°C en el caso de *Pseudomonas* y 180 rpm cuando se trataba de cultivos líquidos.

II.2. Conservación del material biológico.

Para conservar las estirpes a largo plazo, se recogieron células de cultivos saturados, crecidos en medio rico, y se congelaron en criotubos, con glicerol 15% y a -80°C.

II.3. Ensayos de crecimiento en distintas fuentes de azufre.

Para testar la capacidad de *E. coli* con los diferentes plásmidos de crecer en distintas fuentes de azufre, se partió de preinóculos saturados en medio LB. Se lavaron con 3 veces con tampón fosfato (Tampón fosfato 50X: $\text{Na}_2\text{HPO}_4 \times 12\text{H}_2\text{O}$ 17 mM, KH_2PO_4 7,35 mM), y se inocularon a D.O.₆₀₀ 0,1 en medio mínimo M9 con MgSO_4 0,2 mM como fuente

de azufre. Tras 24 horas, se lavaron 3 veces con tampón fosfato y se diluyeron a D.O.₆₀₀ 0,05 ó 0,1 en medio mínimo M9 con la fuente de azufre requerida.

En el caso de las estirpes de *Pseudomonas*, se partió igualmente de preinóculos saturados en medio LB, y se lavaron también 3 veces con tampón fosfato, inoculándose posteriormente en medio mínimo M9/A9 con MgSO₄ 0,2 mM como fuente de azufre. Tras 24 horas, se levaron 3 veces con tampón fosfato y se diluyeron a D.O.₆₀₀ 0,075 en el mismo medio, pero sin azufre. Se dejaron los cultivos toda la noche, para agotar el azufre acumulado, y se lavaron de nuevo 3 veces con tampón fosfato, tras lo cual se diluyeron a 0,05 en medio M9/A9, con las distintas fuentes de azufre necesarias para el experimento.

II.4. Ensayos de toxicidad de 2-hidroxibifenilo (2-HBF).

Se realizaron 2 tipos de ensayos de toxicidad de 2-HBF:

II.4.1. Ensayo de toxicidad de 2-HBF en cultivo líquido.

Se empleó la estirpe MPO554Nal^R, con el fósido UPO112. El ensayo se realizó tal y como se describe en el apartado II.3 de materiales y métodos, pero adicionando al medio distintas concentraciones de 2-hidroxibifenilo (disuelto en NN-DMF)

II.4.2. Ensayo de toxicidad de 2-HBF en medio sólido.

Se realizaron ensayos de supervivencia, partiendo de cultivos en fase exponencial tardía, con diluciones seriadas y siembra en gotas.

Para la estirpe MPO554Nal^R, se empleó en el medio M9 descrito en II.1 de esta sección, pero adicionando al medio distintas concentraciones de 2-hidroxibifenilo (disuelto en NN-DMF) y con MgSO₄ 1 mM como fuente de azufre. Para las estirpes de *Pseudomonas*, se empleó la variante del medio (Mandelbaum *et al.*, 1993), descrita en II.1, adicionando distintas concentraciones de 2-hidroxibifenilo.

II.5. Transferencia de ADN a estirpes bacterianas:

II.5.1. Transformación.

II.5.1.a. Transformación por choque térmico de *E. coli*.

II.5.1.a.1. Preparación de células competentes.

La preparación de células competentes se realizó mediante dos procedimientos distintos dependiendo de la finalidad de la transformación.

Método 1, para la transformación de *E. coli* DH5α con mezclas de ligación:

Se prepararon las células competentes según una variante del método de Inoue *et al.*, 1990, con el cual se obtiene una frecuencia de hasta 10⁹ transformantes por µg

de ADN plasmídico. Para ello un cultivo saturado de DH5 α se diluyó 100 veces en 200 ml de medio SOC (20 g/l de triptona, 5 g/l de extracto de levadura, 0,5 g/l de NaCl, 1,25 mM de KCl; a pH 7, con 10 mM de MgCl₂ y 2,5 g/L de glucosa añadidos en el momento) y se incubó con agitación a 22 °C hasta alcanzar una D.O.₆₀₀ de 0,5. El cultivo se enfrió rápidamente en hielo, y se mantuvo durante 10 minutos. Las células se recogieron mediante centrifugación a 2500 g a 4 °C durante 10 minutos. El sedimento se resuspendió en 20 ml de TB (Pipes (ácido libre) 10 mM, CaCl₂·2H₂O 15 mM, KCl 250 mM, pH a 6.7. y MnCl₂·4H₂O 55 mM y esterilizar por filtración) a 0 °C, se añadieron otros 60 ml de TB a 0 °C y se incubó 10 minutos en hielo. Se recogieron las células en las mismas condiciones y se resuspendieron en 20 ml de TB a 0 °C. Se añadieron 1,5 ml de DMSO y, tras incubación 10 minutos en hielo, se prepararon alícuotas de 0,1 ml, se congelaron en nitrógeno líquido y se almacenaron a -80 °C.

Método 2: Para la transformación con plásmidos de estirpes de *E. coli*:

En este caso las células se hicieron competentes mediante el tratamiento con TSS (Chung *et al.*, 1989). El cultivo saturado de la estirpe receptora se diluyó 100 veces en medio LB y se incubó a 37 °C en agitación hasta fase exponencial temprana. Se repartieron alícuotas de 1 ml en tubos Eppendorf. Se enfriaron en hielo 5 minutos y se centrifugaron 30 segundos en una microcentrífuga. Se desechó el sobrenadante y se añadieron 75 μ l de LB frío. Se incubó en hielo 5 minutos, se añadieron 75 μ l de TSS 2X frío (triptona 0,8 % (p/v), extracto de levadura 0,5 % (p/v), NaCl 85,5 mM, PEG-8000 25 mM, DMSO 10% (v/v), MgSO₄ 100 mM, pH 6,5), se mezclaron suavemente y se incubaron de nuevo 5 min a 4 °C. Las células se transformaron inmediatamente después.

II.5.1.a.2. Transformación.

Para transformar las células competentes se añadieron 10 μ l de mezcla de ligación a 100 μ l de células competentes (del método 1) ó de 1 a 5 μ l de preparación de plásmido (unos 200 ng de ADN), a 150 μ l de células competentes (del método 2). La mezcla se incubó en hielo durante 30 minutos y después se le dio un choque térmico a 42°C, durante 40 segundos, transfiriéndose inmediatamente después, de nuevo a hielo. Finalmente se reconstituyeron las células con 0,9 ml de SOC, se incubaron 90 minutos a 37°C con agitación y se sembraron en placas de medio selectivo.

II.5.1.b. Electrotransformación.

II.5.1.b.1. Preparación de células electrocompetentes.

En el caso de necesitar transformar estirpes de *E. coli*, diferentes a DH5 α , o de las derivadas de EPI300T1^R, o necesitar transferir plásmidos para los que se requirieran

altas frecuencias de transformación, se prepararon células electrocompetentes. Para la preparación de dichas células, se diluyó un cultivo saturado a D.O.₆₀₀ de 0,1 en medio LB y se incubó a 37 °C en agitación hasta D.O.₆₀₀ de 0,6-0,8. Se tomaron alícuotas de 10 ml de células y se recogieron por centrifugación a 16.100 g durante 1 minuto a 4 °C. A continuación, se resuspendieron en 5 ml de glicerol 10 % en agua bidestilada (v/v) frío y se centrifugaron a 16.100 g durante 1 minuto a 4 °C. Seguidamente se repitió el proceso resuspendiendo las células en 1 ml de glicerol 10 % frío, y nuevamente en 0,5 ml de glicerol 10 % frío. Finalmente se centrifugaron de nuevo y se resuspendieron en 0,1 ml de glicerol 10% frío, se incubaron en hielo y se transformaron inmediatamente.

II.5.1.b.2. Electroporación.

Para transformar las células electrocompetentes se añadió de 1 a 5 µl de preparación de plásmido (unos 200 ng de ADN), ó 10µl de mezcla de ligación, a 100 µl de células. La mezcla se transfirió a una cubeta de electroporación con una distancia entre los electrodos de 2 mm, previamente enfriada, y se sometió a una diferencia de potencial de 2,5 KV en un electroporador MicropulserTM (Bio-Rad). Inmediatamente después del pulso, se añadió 1 ml de LB a temperatura ambiente a las células. Se incubaron a 37 °C durante 90 minutos, en agitación y se sembraron en placas de medio selectivo.

II.5.2. Conjugación.

II.5.2.a. Conjugación triparental.

En el caso de necesitar transferir plásmidos movilizables, con un origen de transferencia conjugativa pero sin genes *tra*, desde una estirpe a otra de *E. coli*, se realizaron conjugaciones triparentales, empleándose los plásmidos auxiliares pRK2013 o pRK2073 (Figurski & Helsinki, 1979). Se emplearon cultivos saturados de las estirpes donadoras, auxiliar y receptora, crecidos en LB. Se diluyeron en medio LB a D.O.₆₀₀ de 0,1 y se incubaron a 37°C hasta D.O.₆₀₀ aproximada de 0,5. Se mezclaron en proporción 1:1:1, en volumen final de 1,5 ml. Se lavó la mezcla con tampón fosfato (Tampón fosfato 50X: Na₂HPO₄ x 12H₂O 17 mM, KH₂PO₄ 7,35 mM), se centrifugó y se eliminó la mayor parte del sobrenadante. El sedimento de células se resuspendió en 50µl de tampón fosfato y las células resuspendidas se depositaron en forma de gota (parche de conjugación) en una placa de LB sin antibióticos. El parche se dejó secar en campana de flujo laminar y se incubó a 37°C durante al menos 8h. Después, se segregó el parche en placa de medio selectivo.

II.6. Rastreo de actividad funcional. Búsqueda de clones con capacidad de crecimiento en DBT.

Para encontrar un clon en la metagenoteca de las pilas de la refinería de la industria petroquímica, que le confiera a *E. coli* la capacidad de crecer con DBT como única fuente de azufre, se realizó una conjugación triparental, con algunas modificaciones, ya que dicha metagenoteca se encuentra almacenada en EPI300T1^R y es necesario transferirla a MPO554, para facilitar los rastreos funcionales.

Se tomó una alícuota de 0,5 ml de la metagenoteca, almacenada a -80°C y se cultivó en un volumen total de 5ml de LB con Cm 12,5 µg/ml, hasta D.O.₆₀₀ de 0,5. Luego se diluyó este cultivo (donador) y cultivos saturados de las estirpes receptora (MPO554 Nal^R) y auxiliar en LB a D.O.₆₀₀ de 0,1. Se incubaron a 37°C hasta D.O.₆₀₀ aproximada de 0,5 y se mezclaron los cultivos en proporción 1:1:1 en un volumen final de 3 ml. Se lavó y se puso el parche de conjugación. Tras la incubación, el parche de conjugación se recogió, se resuspendió en tampón fosfato (Tampón fosfato 50X: Na₂HPO₄ x 12H₂O 17 mM, KH₂PO₄ 7,35 mM) se hizo una dilución 1/10 en 1 ml de tampón fosfato y se sembraron 10 placas en medio M9 con DBT como única fuente de azufre, arabinosa y salicilato. Para calcular la frecuencia de conjugación, se realizaron diluciones del parche resuspendido en 2 ml de tampón y se sembraron diluciones de 10⁴ y 10⁵ en placas de LB con Cm 12,5 µg/ml (resistencia de la metagenoteca) y Nal 15µg/ml (resistencia de la estirpe receptora) para el conteo de transconjugantes y por otro parte se sembraron diluciones de 10⁶ y 10⁷ en placas de LB Nal 15µg/ml, para el conteo de viables. Tras la incubación, se realizaron conteos de las colonias obtenidas y se calculó la frecuencia de conjugación dividiendo el número de transconjugantes entre el de viables.

III. Técnicas bioquímicas:

III.1. Ensayos de actividad β-galactosidasa.

La inducción de las fusiones génicas *dsz::lacZ*, para su posterior ensayo de actividad β-galactosidasa se realizó de diferente manera según los experimentos:

En el caso de las fusiones a *dszE_{UPO21}*, *dszEA_{UPO21}* y *dszEABC_{UPO21}* se partió de preinóculos saturados en medio LB. Se lavaron con tampón fosfato (Tampón fosfato 50X: Na₂HPO₄ x 12H₂O 17 mM, KH₂PO₄ 7,35 mM) y se inocularon a D.O.₆₀₀ 0,05 en medio mínimo M9, con MgSO₄ 2 mM como fuente de azufre. Tras 24 horas, se lavaron 3 veces con tampón fosfato y se diluyeron a D.O.₆₀₀ 0,1 en medio mínimo M9 con la fuente de azufre requerida, y los inductores necesarios en cada caso, dejándolos crecer entre 6 y 24 horas, según el experimento, a 30°C. (En el caso de que las fusiones se

quisieran inducir con DBT, se aportaba también 0,2 mM de SO_4 , ya que no pueden crecer en DBT como única fuente de azufre).

En cuanto a la fusión a *dszE_{UPO112}*, se partió de preinóculos saturados en medio LB. Se lavaron con tampón fosfato y se inocularon a D.O.₆₀₀ 0,1 en medio mínimo M9 con MgSO_4 0,2 mM como fuente de azufre. Tras 24 horas, se lavaron 3 veces con tampón fosfato y se diluyeron a D.O.₆₀₀ 0,1 en medio mínimo M9 con la fuente de azufre requerida y los inductores necesarios en cada caso, dejándolos crecer 24 horas, a 30°C. En el caso de que los experimentos requirieran el aumento del nº de copias (inducción con arabinosa), ésta se adicionó también en los preinóculos.

Tras el período de incubación, se analizó la expresión génica siguiendo el método descrito por Miller (Miller, 1992). Se tomaron 100 µl de cultivo, diluido a D.O.₆₀₀ 0,15, y se añadieron a una mezcla de 650 µl de Tampón Z (Na_2HPO_4 60 mM, $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ 40 mM, KCl 10 mM, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 1 mM), (con 0,27 % (v/v) de β-mercaptoetanol recién añadido), 20 µl de SDS 0,1% (p/v) y 30 µl de cloroformo. La mezcla se agitó en el vórtex para permeabilizar las células, y se atemperó 10 min a 30 °C. Después, a cada muestra se le añadieron 200 µl de solución de ONPG 13 mM, para iniciar la reacción y se continuó incubando a 30°C hasta observar el color amarillo debido al ONPG en las muestras. En ese momento, se detuvo la reacción adicionando 500 µl de Na_2CO_3 1M, homogeneizando con vórtex y enfriando en hielo. Se midió el tiempo de reacción, es decir, el tiempo transcurrido entre la adición de ONPG y la de Na_2CO_3 , se centrifugó la mezcla 15 min a 9300 g y temperatura ambiente, y se midió la absorbancia de la fase acuosa a 420 nm.

La actividad se calculó utilizando la siguiente fórmula:

$$\text{Actividad} = (A_{420} : (V \cdot \text{D.O.}_{600} \cdot t)) \cdot 1000$$

Actividad: Unidades Miller.

A_{420} : Absorbancia de la reacción a 420 nm de longitud de onda.

V: Volumen de cultivo utilizado en la reacción (ml).

D.O.₆₀₀: Densidad óptica del cultivo a 600 nm de longitud de onda.

t: Tiempo de reacción (minutos).

En todos los casos se realizaron al menos tres réplicas independientes de cada experimento, calculándose los valores medios y desviaciones estándar entre las réplicas.

III.2. Manipulación de proteínas.

III.2.1. Electroforesis de proteínas en geles de poliacrilamida-SDS.

Para analizar la cantidad total de proteína de la célula en geles, se centrifugó 1 ml de cultivo, y se resuspendió en 25 μ l de agua bidestilada, por cada unidad de DO_{600} del cultivo, y se le añadió un volumen de tampón de carga 2X (Tris-HCl 160 mM pH 6,8, glicerol 20 % (v/v), SDS 4 % (p/v), azul de bromofenol 0,1 % (p/v), β -mercaptoetanol 10 % (v/v)). Se hirvieron las muestras 5 minutos, tras lo cual se centrifugaron, y se cargó en gel el sobrenadante (normalmente 5 μ l). En el caso de proteínas solubles, se diluyó la proteína en tampón de lisis, se le añadió un volumen de tampón de carga 2X, se hirvieron las muestras 5 minutos, y se cargaron en gel.

Las proteínas se analizaron mediante electroforesis en geles verticales desnaturalizantes de poliacrilamida con SDS 0,1% (p/v) de acuerdo con el método descrito por Laemmli (Laemmli, 1970). La electroforesis se realizó en geles discontinuos, compuestos por un gel superior de empaquetamiento (Tris-HCl 125 mM pH 6,8, SDS 0,1% (p/v), acrilamida:bisacrilamida (36,5:1) 4% (p/v), persulfato amónico 0,1% (p/v), TEMED 0,1% (p/v)) y un gel inferior de separación (Tris-HCl 375 mM pH8,8, SDS 0,075% (p/v), acrilamida:bisacrilamida (36,5:1) entre 7,5 y 15% (p/v) según el tamaño de la proteína a analizar, persulfato amónico 0,033% (p/v), TEMED 0,067% (p/v)).

Los geles se corrieron en un sistema vertical Mini Protean® III (Bio-Rad) a una intensidad de corriente de 30-35 mA/gel en Tank buffer 1X (Trizma base 25 mM, glicina 19,2 mM y SDS 0,1% p/v) usando como marcador de peso molecular el marcador B (Bio-Rad), compuesto por una mezcla de nueve proteínas de entre 6,5 y 200 KDa. Posteriormente se lavaron con agua bidestilada y se tiñeron con el reactivo EZ-Blue Gel staining reagent (Sigma) siguiendo las instrucciones del fabricante, tras lo cual, se lavaron con agua, y finalmente, se secaron en un secador Hoefer™ Slab Gel Dryer CD 2000 (GE healthcare) colocando los mismos sobre dos láminas de papel Whatman 3MM y sometidos durante 60 min a vacío y 80 °C.

III.2.2. Superproducción de DszR.

Para superproducir la proteína se utilizó la estirpe NCM631 transformada previamente con el plásmido pLZ227, que contiene un sistema de expresión basado en la ARN polimerasa de T7 (Govantes & Santero, 1996), en la cual se introdujo el plásmido pMPO1554, que contiene *dszR* con su extremo amino fusionado a una inteína. La estirpe NCM631 con ambos plásmidos, se creció en LB con Ap y Cm hasta saturación, y se diluyó a $D.O_{600}$ de 0,1 en 1 l de medio fresco, incubándose a 37°C hasta $D.O_{600}$ 0,3. El crecimiento se detuvo poniendo el cultivo en hielo, y se indujo la expresión del promotor de T7 con IPTG 0,4 mM, tras lo cual, se continuó la incubación durante 21,5 h

a 16°C. Finalmente, las células se recogieron por centrifugación a 4000 g y 4°C durante 20 minutos y el pellet se conservó a -80°C, para su posterior utilización.

III.2.3. Purificación de DszR

Para la purificación de DszR_{UP021}, se utilizó el sistema IMPACT™ de New England Biolabs (NEB E6901), con algunas modificaciones. Todo el proceso se realizó a 4°C.

El pellet de la superproducción, se resuspendió en tampón de columna (tampón Tris-HCl 20 mM, pH 8, con NaCl 100 mM y EDTA 1 mM) al que se añadió protease inhibitor cocktail (1183617001 Roche). Tras lo cual, se rompió por sonicación. (Sonifier S-250D, BRANDSON). La rotura se realizó en modo manual con 30% de amplitud, durante 5 minutos totales con pulsos de 2 segundos on, 2 segundos off. Tras la rotura, las células se centrifugaron durante 30 minutos a 15000g a 4°C y el extracto clarificado se cargó en una columna de quitina, preequilibrada con el tampón de columna. La columna se lavó abundantemente con el mismo tampón y se indujo el corte con: Tris-HCl 20 mM, pH 8, NaCl 0,5 M, EDTA 1 mM y DTT 50 mM. La rotura de la proteína de fusión se realizó durante 72h a 18°C.

DszR se eluyó y se dializó en un casete de diálisis para reducir las concentraciones de NaCl, EDTA y DTT. La diálisis se realizó en 2 pasos; en primer lugar, se mantuvo 18 horas a 4 °C, con tampón Tris-HCl 20 mM, pH 8, NaCl 0,3 M, EDTA 0,2 mM y 10 mM DTT. El segundo paso, se realizó a 18°C, durante 3 horas con tampón Tris-HCl 20 mM, pH 8, NaCl 0,2 M, EDTA 0,2 mM y 5 mM DTT. Finalmente, la proteína se distribuyó en pequeñas alícuotas y se almacenó a -80°C, para su posterior utilización. Su pureza se estimó visualmente en gel.

III.2.4. Cuantificación de proteínas.

DszR_{UP021} se cuantificó con el kit comercial RC/DC de Bio-Rad (5000121) siguiendo las instrucciones del fabricante.

III.3. Cuantificación de la producción de 2-hidroxibifenilo.

La determinación de la cantidad de 2-HBF producido, se realizó por el método de Gibbs. Las distintas estirpes se crecieron en sus correspondientes medios mínimos, con DBT como fuente de azufre, se tomaron muestras a diferentes tiempos y se ensayaron según lo descrito por Murarka (Murarka *et al.*, 2019). Se utilizaron muestras del cultivo de 10 ml, (y una muestra de medio mínimo sin azufre y sin inocular para el blanco) y se ajustó su pH a 8. Se diluyeron entre 2 y 5 veces en 1 ml de volumen final y se añadió 10 µl de reactivo de Gibbs (Sigma D6511) al 1% (p/v) en etanol. Se incubaron

durante 30 minutos en oscuridad a 30°C, tras lo cual se centrifugaron 2 minutos a velocidad máxima y se midió la D.O.₆₁₀ del sobrenadante.

Para calcular la cantidad de 2-HBF producida se realizó una recta patrón con este compuesto, en las mismas condiciones.

IV. Técnicas de biología molecular:

IV.1. Manipulación de ácidos nucleicos.

IV.1.1. Manipulación de ADN.

IV.1.1.a. Aislamiento de ADN plasmídico de *Escherichia coli*.

Las preparaciones de ADN plasmídico de las distintas estirpes de *E. coli*, se realizaron mediante 3 métodos diferentes, según las necesidades del experimento.

Para obtener ADN de alta pureza, necesaria para su posterior secuenciación, se utilizó el kit Nucleospin Plasmid Mini (Macherey-Nagel) siguiendo las instrucciones del fabricante.

Para obtener ADN de pureza superior, necesario para servir como molde en transcripciones *in vitro*, se empleó el kit Midiprep Nucleobond PC 100 (Macherey-Nagel), siguiendo las instrucciones del fabricante para la purificación de plásmidos de bajo número de copia.

Para el resto de los usos habituales en el laboratorio, la extracción de ADN se realizó siguiendo el protocolo simplificado de lisis alcalina (Stephen *et al.*, 1990). Para lo cual, se recogieron 3-5 ml de cultivo saturado, centrifugando a máxima velocidad durante 30 segundos. Después, se resuspendió el sedimento en 100 µl de solución I (glucosa 50 mM, EDTA 10 mM, Tris-HCl 25 mM, pH 8) y se incubó 5 minutos a temperatura ambiente, tras lo cual se añadieron 200 µl de solución II (NaOH 0,2 N, SDS 1 % (p/v) preparada en el momento), se agitaron los tubos por inversión y se mantuvieron en hielo durante 5 minutos. Tras esta incubación, se añadieron 150 µl de solución III (acetato potásico 3 M ajustado a pH 4,8 con ácido fórmico) y se mantuvieron los tubos en hielo, invirtiéndolos repetidas veces durante 5 min. Posteriormente se centrifugaron a máxima velocidad durante 10 minutos, y el sobrenadante se transfirió a un tubo nuevo, al cual se le añadió un volumen de etanol 96 % (v/v) enfriado a -20°C y se centrifugó inmediatamente durante 5 minutos a máxima velocidad. Seguidamente, se lavó el precipitado con 1 ml de Etanol 70 % (v/v) enfriado a -20°C, se centrifugó 3 minutos a máxima velocidad y se secó mediante vacío. Finalmente, se resuspendió en 25-50 µl (según el número de copias del plásmido), de TER (Tris-HCl 10 mM pH 8, EDTA 1 mM, ribonucleasa 0,02 % (p/v)) y se incubó 15 minutos a 37°C.

IV.1.1.b. Electroforesis de ADN.

Para visualizar fragmentos lineales o preparaciones de ADN, se realizaron geles de agarosa de baja EEO, cuya concentración osciló entre 0,6 y 1,5 % (p/v) en tampón TAE (Tris 40 mM, ácido acético 3,4 M, pH 7,7 y EDTA 10 mM) según el tamaño de los fragmentos a visualizar. Se han utilizado 3 patrones de peso molecular: marcador 1 kb Plus Ladder (Invitrogen) y GeneRuler 1 kb Plus (Thermo Fisher) para muestras de hasta 12 kb, y para muestras superiores se ha empleado ADN del fago lambda digerido con *HindIII*. Para cargar las muestras en gel, se les añadió tampón de carga 6X (azul de bromofenol 0,05 % (p/v), xilen cianol 0,05 % (p/v), glicerol 30 % (v/v), Tris-HCl 10 mM pH 8 y EDTA 1 mM pH 8). Los geles se sometieron a una diferencia de potencial de entre 40 y 130 V.

Tras la electroforesis, los geles se tiñeron con una solución de bromuro de etidio a 1 mg/l en agua destilada. Las bandas de ADN se visualizaron con un transiluminador de radiación ultravioleta (Enduro GDS) y se registraron digitalmente, con el mismo dispositivo. En los casos en los que se pretendía recuperar ADN de gel se minimizó la exposición a radiación ultravioleta para evitar daños en el ADN.

IV.1.1.c. Reacción en cadena de la polimerasa (PCR).

Cuando fue necesario amplificar fragmentos de ADN para su posterior clonación, se realizaron reacciones de PCR utilizando el ADN correspondiente como molde. Las reacciones se llevaron a cabo en un volumen final de 25 μ l en tubos de 0,2 ml, en presencia de los oligonucleótidos adecuados a una concentración de 1 μ M, dNTPs a 100 μ M y 1 unidad de ADN polimerasa termorresistente de alta fidelidad (Q5 New England Biolabs), en el tampón que suministra la casa comercial, que contiene $MgCl_2$ a una concentración final de 1,5 mM. Cuando fue necesario se varió la concentración de Mg^{+2} . Como molde se utilizaron: 50 ng de ADN cromosómico, 20 ng de ADN plasmídico, o ADN procedente de células hervidas. Para ello se resuspendió una colonia en 50 μ l de H_2O bidestilada autoclavada y se hirvió a 100 °C durante 5 min. Pasado este tiempo, la mezcla se centrifugó 30 s a velocidad máxima y se usaron 5 μ l del sobrenadante como molde para la reacción de amplificación. Las reacciones se llevaron a cabo en un termociclador Biometra Tpersonal.

El programa de PCR del termociclador varió en función del tamaño y naturaleza del fragmento a amplificar, pero por lo general consistió en:

a) Desnaturalización inicial: 2-5 min a 94 °C

b) 35 ciclos: 30 s a 94 °C (desnaturalización), 30 s a la temperatura específica de desnaturalización de los oligonucleótidos (T_m) (apareamiento) y entre 30-60 s/kb de ADN a amplificar a 72 °C (extensión).

c) Extensión final: 5 min a 72 °C.

Los oligonucleótidos utilizados durante esta Tesis fueron sintetizados por la empresa Sigma y sus características principales se detallan en la Tabla 3.

IV.1.1.d. Aislamiento y purificación de fragmentos de ADN.

Los fragmentos de ADN a purificar, se corrieron en gel de agarosa y se identificaron mediante visualización con un transiluminador de radiación ultravioleta de baja energía, posteriormente, se aislaron recortándolos del gel. Para purificarlos, se empleó el kit GFX (GE Healthcare) y se siguieron las instrucciones del fabricante.

IV.1.1.e. Cuantificación de preparaciones de ADN.

Para cuantificar las diferentes preparaciones de ADN, se empleó un espectrofotómetro Nanodrop ND-100 Spectrophotometer (Thermo Scientific), que midió la absorbancia a 260 nm y calculó la concentración, aplicando un coeficiente de extinción de 50 ng cm μl^{-1} .

IV.1.1.f. Clonación de fragmentos de ADN.

Para la clonación de fragmentos de ADN en vectores, se digirió el plásmido o el producto de PCR con las enzimas de restricción adecuadas en cada caso (procedentes de las compañías Roche o New England Biolabs), empleando el tampón comercial correspondiente, según las instrucciones del fabricante. En el caso de requerir la conversión de extremos cohesivos en romos, si se trataba de rellenar los extremos 3' de los fragmentos 5' protuberantes se empleó ADN polimerasa Klenow (Roche) o ADN polimerasa de T4 (Roche), y si había que eliminar algún extremo 3' protuberante se empleó ADN polimerasa de T4 (Roche). En algunos casos fue necesaria la desfosforilación de los extremos de los vectores para evitar religaciones en las clonaciones, para lo que se empleó la enzima SAP (Shrimp Alkaline Phosphatase, USB) siguiendo las instrucciones del fabricante.

Las reacciones de ligación se realizaron en un volumen final de 10 μl , con la enzima ligasa (M0202) New England Biolabs, según las instrucciones del fabricante. La mezcla de ligación se transformó en la estirpe *E. coli* DH5 α competente tal y como se explica en el apartado II.5.1.a de materiales y métodos. Los clones obtenidos se comprobaron mediante análisis de restricción y posterior secuenciación, si el fragmento clonado provenía de PCR.

IV.1.1.g. Secuenciación y análisis de secuencias.

En las construcciones de los plásmidos realizadas por PCR, se confirmó la presencia o ausencia de mutaciones en el ADN, según el caso, mediante reacciones de secuenciación, que fueron llevadas a cabo por las empresas Secugen (Madrid) o StabVida (Caparica, Portugal) utilizando como molde preparaciones de ADN plasmídico o productos de PCR de alta calidad. El análisis de las secuencias obtenidas se realizó empleando el paquete de herramientas BLAST online (Altschul *et al.*, 1997) disponible en la página web del National Center for Biotechnology Information (NCBI) (<http://blast.ncbi.nlm.nih.gov>).

Las secuencias de ADN se tradujeron a proteínas con la herramienta ExPaSy Translate, tool (<https://web.expasy.org/translate/>).

La secuenciación del fósido UPO112 se llevó a cabo con un secuenciador Roche 454 GS FLX Ti (454 Life Sciences, Branford, CT, USA) por la empresa Lifesequencing (Valencia), a partir de preparaciones de ADN de alta calidad. El ensamblaje de las secuencias se realizó mediante Newbler GS De Novo Assembler v.2.3 (Roche). Se realizaron comparaciones con las de las bases de datos nucleotide collection, non-redundant protein sequences y UniProtKB/Swiss-Prot empleando el paquete de herramientas BLAST.

Para los alineamientos de secuencias se emplearon las herramientas BLAST y Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

Para la construcción de árboles filogenéticos, el alineamiento múltiple de secuencias se realizó con el programa MEGAX (Kumar *et al.*, 2018) utilizando el método MUSCLE. Los árboles se estimaron usando el método de Máxima Verosimilitud (*Maximum Likelihood*), eligiendo el mejor modelo indicado por MEGAX en cada caso (que fue WAG+G+F para el árbol de DszE y LG+G para el árbol de DszR) y seleccionando en el parámetro *Gap/Missing Data Treatment* la opción *Partial Deletion*. Se realizó un análisis bootstrap, indicando un número de réplicas de 100.

Las estructuras secundarias del ARN se estudiaron con la herramienta: RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>)

La traducibilidad de las regiones de interés se analizó con la herramienta RBS calculator (<https://www.denovodna.com/software/reverse>) y el método “reverse Engineer RBSs method” (Salis *et al.*, 2009).

IV.1.2. Manipulación de ARN.

IV.1.2.a. Extracción de ARN total de *Escherichia coli*.

Las extracciones de ARN se realizaron a partir de distintos volúmenes de cultivo en función del experimento a realizar. Cuando los cultivos alcanzaron la fase exponencial temprana, las células se centrifugaron a máxima velocidad y 4°C en alícuotas de 2 ml, se eliminó el sobrenadante y el sedimento se congeló inmediatamente en nitrógeno líquido. Las alícuotas se guardaron a -80 °C hasta el momento de ser procesadas. Posteriormente, las alícuotas que venían de un mismo cultivo se resuspendieron en 0,5 ml de TriSure Isolation Reagent (Bioline), que contiene fenol e isocianato de guanidinio, se mezcló por inversión y se incubó 10 min a 60 °C para lisar las células. Se centrifugaron 10 min a 16100 g y a 4 °C para eliminar los restos celulares y se transfirieron a tubos de Phase Lock Gel (5 prime) precompactados. A cada tubo, se le añadieron 200 µl de cloroformo, se mezclaron por inversión y se incubaron a 4 °C durante 15 min. Después se centrifugaron 15 min a 16100 g y a temperatura ambiente y se añadieron 400 µl de fenol:cloroformo:isoamílico (25:24:1), se mezclaron de nuevo por inversión, y se centrifugaron 5 min a 16100 g y a 4 °C. Se transfirió la fase acuosa a otro tubo, y se precipitó con 1 volumen de isopropanol, incubándolo durante 10 min a 4 °C. Tras lo cual, se centrifugaron 10 minutos a 4°C y máxima velocidad. Posteriormente, el sedimento se lavó con etanol 70% preenfriado, y se centrifugó 5 minutos, a 4 °C y 16100 g. Después, se eliminó el sobrenadante y se secaron los precipitados en campana de flujo laminar. Finalmente, se resuspendieron en un volumen de 50 µl de H₂O DEPC (preparada incubando a 37°C, 1 ml de DEPC en 1 l de agua bidestilada, durante al menos 12 h y autoclavándolo posteriormente 2 veces).

Para las reacciones de extensión desde el cebador (*primer extension*), el ARN se cuantificó midiendo la absorbancia a 260 nm en un espectrofotómetro NanoDrop ND-1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) se visualizó en gel de agarosa para verificar la exactitud de la cuantificación y comprobar su integridad, y se congeló a -80 °C en alícuotas de 20 µg.

Para los ensayos de PCR cuantitativa a las muestras de ARN se les realizó un tratamiento con DNAsa para eliminar posibles trazas de ADN utilizando para ello el kit DNA free (Ambion) siguiendo las instrucciones del proveedor. Tras lo cual se limpiaron con el kit RNeasy mini kit (Qiagen), usando el protocolo Clean Up, eliminando gran parte del ARN ribosómico 5S y eluyendo cada muestra en 50 µl de H₂O DEPC.

Para comprobar la ausencia de contaminación con ADN se llevó a cabo una PCR, utilizando 500 ng de ARN como molde y oligonucleótidos específicos de la estirpe en cuestión, y se comprobó la no aparición de producto de amplificación mediante electroforesis. En el caso de que hubiese ADN contaminante, se repitió el tratamiento

con DNAsa una o más veces si fuese necesario. Las muestras de ARN fueron guardadas a -80 °C hasta el momento de su uso.

IV.1.2.b. Electroforesis de ARN.

Para visualizar y comprobar la integridad de las muestras de ARN, éstas se corrieron en geles de agarosa al 1% (p/v) (según lo descrito en materiales y métodos IV.1.1.b), que contenían 0,05 mg de bromuro de etidio por litro de agarosa. El gel se visualizó con un transiluminador de radiación ultravioleta (Enduro GDS) y se registraron con un analizador de imágenes.

IV.1.2.c. Cuantificación de ARN.

Las preparaciones de ARN se cuantificaron midiendo la absorbancia a 260 nm, en un espectrofotómetro Nanodrop ND-100 Spectrophotometer (Thermo Scientific), que calculó la concentración usando un coeficiente de extinción de 40 ng cm μl^{-1} .

IV.1.2.d. Retrotranscripción de ARN

Se mezclaron cantidades equivalentes de ARN de 3 réplicas técnicas a una concentración final de ≈ 300 ng/ μl . Se emplearon 10 μl de dicha mezcla para las reacciones de retrotranscripción, que se realizaron con el kit High Capacity cDNA Reverse Transcription (Applied Biosystems) siguiendo las instrucciones del fabricante. Acabada la reacción, se limpiaron con el kit qiaquick PCR (qiagen) y se eluyeron en 50 μl de H₂O tratada con DEPC.

IV.2. Determinación del extremo 5' de ARN mensajeros por el método de extensión desde el cebador (*primer extension*)

La determinación del inicio de la transcripción de los transcritos se realizó según lo descrito por Govantes *et al.*, 2000, con algunas modificaciones.

IV.2.1. Marcaje de los oligonucleótidos con $\gamma[^{32}\text{P}]\text{ATP}$

Se diseñó un cebador de 30 nucleótidos, complementario a la cadena codificante del gen objeto de estudio, llamado PEX-DBT1. Dicho cebador se marcó radiactivamente en el extremo 5' con $\gamma[^{32}\text{P}]\text{ATP}$, utilizando la enzima polinucleótido quinasa de T4 (Roche). Esta reacción de fosforilación se realizó en un volumen de 10 μl mezclando 5 pmoles del oligonucleótido (1 μl oligo 5 μM), 70 μCi de $\gamma[^{32}\text{P}]\text{ATP}$, tres unidades de polinucleótido quinasa T4 y el tampón de la casa comercial. Se incubó durante 1 hora a 37 °C y se paró con 40 μl de EDTA 25 mM. Los nucleótidos y la marca no incorporados,

se eliminaron con una columna de sephadex G25 (GE Healthcare), tras lo cual, se fenolizó el cebador marcado, con fenol:cloroformo 1:1 en un tubo de Phase Lock Gel (Eppendorf) precompactado.

La cantidad de marca se cuantificó en un contador de centelleo.

IV.2.2. Reacción de extensión de la retrotranscriptasa.

Para llevar a cabo la extensión del cebador, se mezclaron 20 µg de ARN total (de las estirpes y condiciones de cultivo requeridas según el experimento), con 300000 cpm del correspondiente oligonucleótido marcado, en un volumen de 50 µl. La mezcla se precipitó a -20°C con 1/10 de volumen de acetato sódico 3 M pH 5,2 y 2,5 volúmenes de etanol 96% (v/v) en presencia de 1µl de glucógeno 20 mg/ml. Tras lo cual, se centrifugó a máxima velocidad durante 20 min a 4 °C. El precipitado se lavó con 120 µl de etanol 80 % (v/v), se secó en campana de extracción, y se resuspendió en 15 µl de tampón de hibridación (10 mM de Tris-HCl (pH 8,3) y 150 mM de KCl). A continuación, se realizaron los pasos de desnaturalización, (5 min a 95 °C) e hibridación (90 min a 60°C) en un termociclador. Tras la hibridación, se añadieron a cada muestra 15 µl de mezcla de extensión (Tris HCl 90 mM (pH 8,3), MgCl₂ 6 mM, DTT 2 mM, dNTPs 2 mM, 4 U de inhibidor de ARNsas (ARNse OUT, Invitrogen) y 100 U de la retrotranscriptasa Superscript III (Invitrogen)) precalentada a 50 °C, y se incubó durante 1 h a 50 °C. Después, se paró la reacción y se degradó el ARN, añadiendo 3 µl de EDTA 250 mM y 3,2 µl de NaOH 2 M (preparada en el momento), incubando 30 min a 65 °C y añadiendo después, 3,2 µl de HCl 2M (preparado fresco). Finalmente, se precipitó a -20°C con 1 µl de glucógeno 20 mg/ml, 10 µl de acetato sódico 3 M (pH 5,2) y 150 µl de etanol 96% (v/v). Tras la precipitación, se centrifugó 20 min a 4°C y máxima velocidad, se lavó y secó el precipitado, y se resuspendió en 5-10 µl de solución de carga (mezcla 3:2 de solución de parada (formamida 95%, EDTA 20 mM, azul de bromofenol 0,05% y xylene cyanol 0,05%) y H₂O).

Para poder localizar el extremo 5' del fragmento retrotranscrito, se realizaron reacciones de secuenciación con el mismo oligonucleótido usado para la extensión, y utilizando como molde un fragmento de ADN de doble cadena que contenía el posible inicio de transcripción. Dicho fragmento, fue secuenciado mediante el método de terminación con dideoxinucleótidos utilizando el kit Thermo Sequenase Cycle Sequencing Kit (USB), siguiendo las instrucciones del fabricante. El programa de PCR utilizado fue una desnaturalización inicial de 5 min a 95 °C y 50 ciclos de 30 s a 95 °C, 30 s a 60 °C y 60 s a 72 °C.

IV.2.3. Análisis en gel de secuenciación.

Las muestras de la extensión se calentaron 3 min a 95°C y las de la secuenciación 2 min a 75°C, tras lo cual, se transfirieron inmediatamente a hielo. Posteriormente se cargaron 2-3 µl de cada reacción en un gel de secuenciación de TBE/acrilamida 6% (p/v) -urea 8M, precalentado. El gel se corrió en TBE 1X (Tris-ácido bórico 90 mM, EDTA 2 mM (pH 8,3)). Tras la carrera, los geles se secaron en un secador HoeferTM Slab Gel Dryer CD 2000 (GE Healthcare) y se expusieron al menos durante 12 h en una pantalla sensible a radiaciones gamma (Storage Phosphor Screen, GE Healthcare). Finalmente, la pantalla se escaneó con un escáner Typhoon 9410 (GE Healthcare).

Para estudiar la expresión de genes a nivel transcripcional se realizaron ensayos de qPCR, según lo descrito por García-Mauriño *et al.*, 2013. Se diseñaron parejas de oligonucleótidos específicos, en las primeras 100 pb de cada gen a estudiar. Para *dszR* se emplearon los oligonucleótidos: *dszRQ* Fw y Rv y para *dszE*: *dszEQ* Fw2 y Rv2. Se preparó una mezcla en frío con 10 ng de cDNA previamente retrotranscrito, con 0,3mM de cada cebador, 5 µl de Fast Gene 2x (bajo Rox) (Nippon Genetics) y 3,4 µl agua bidestilada. Como molde para la recta de calibrado, se emplearon diluciones seriadas de ADN del clon objeto de estudio (concentraciones entre 25 y 0,0025 ng). Cada reacción se realizó al menos por triplicado. Las reacciones de PCR cuantitativa se realizaron en un termociclador de tiempo real CFX de BioRad.

Para analizar la capacidad de activación de *DszR*_{UPO21} sobre *P*_{dszUPO21}, se llevaron a cabo ensayos de transcripción *in vitro*, (IVT). Los cuales, se realizaron en un volumen final de reacción de 20 µl que contenía 40 mM Tris HCl (pH 7,5), 150 mM KCl, 10 mM MgCl₂, 2% glicerol, 1 mM DTT, 5 mg BSA, 4 mM ATP y 0,5 µg de plásmido molde superenrollado, que contenía *P*_{dszUPO21} (pMPO1548), o el vector vacío (pTE103). Se preparó una mezcla con *core* de la ARN polimerasa de *E. coli*, (Epicentre) (100 nM), σ^{54} de *P. putida* (80 nM), e IHF de *E. coli* (75 nM), y se incubó a 30°C durante 10 minutos, tras lo cual, se añadieron concentraciones crecientes de *DszR* a la mezcla y se incubó a 30°C, durante 10 minutos. *E. coli* IHF fue cedida por Fernando Govantes y Ana Platero (Platero *et al.*, 2016). *P. putida* σ^{54} fue cedida por Victoria Shingler. Después de la incubación, se añadió una mezcla de ATP, GTP, CTP (concentración final de cada uno 0.4 mM), UTP (0.07 mM) y [α -³²P]-UTP (0,033 mM, Perkin Elmer) para

iniciar la transcripción *in vitro* multirronda. Tras 5 minutos a 30°C, se añadió heparina (0,1 mg ml⁻¹). Se incubaron 5 minutos a 30°C y finalmente se paró la reacción con 5 µl de tampón de parada (0,125% p/v azul de bromofenol, 0,125% p/v xylene cyanol, 20 mM EDTA, 95% v/v formamida) y se mantuvieron en hielo. Las muestras se hirvieron durante 2 minutos y se sometieron a electroforesis en geles de 1,5 mm de grosor de TBE/acrilamida 4 % (p/v)-urea 8 M previamente precorridos en TBE 1X. Los geles se secaron con un secador Hoefer™ Slab Gel Dryer CD 2000 (GE Healthcare) y se expusieron a una pantalla sensible a radiaciones gamma (Storage Phosphor Screen, GE Healthcare). Las bandas se visualizaron con un escáner Typhoon 9410 y se analizaron usando el software ImageQuant (GE Healthcare).

Los fragmentos de ADN para el análisis de la interacción con DszR se obtuvieron por restricción del plásmido pMPO1547 que contiene los posibles sitios de unión de DszR. Cuando se necesitó marcar la cadena +, con respecto al inicio de la transcripción de *dszE* (para ensayos de retardo en gel, o de protección frente a digestión por DNasa I), se usaron las restricciones *Sall-EcoRV*, y en el caso de necesitar marcar la cadena - (para ensayos de protección frente a digestión por DNasa I), se emplearon las digestiones *Apal-HindIII*. Las sondas generadas en ambos casos, se marcaron en su extremo 5' con α[³²P]dCTP; mezclando 1 pmol del fragmento de ADN, dGTP, dATP, dTTP y α[³²P]dCTP 250 µM, 1 unidad de Klenow (NEB) y tampón ligasa de la misma casa comercial. La reacción se incubó a 25°C durante 15 minutos, y posteriormente el ADN se purificó con el kit MSB Spin PCRapace kit (Invitex), eluyéndose en 50 µl de agua bidestilada, para obtener una concentración aproximada de sonda de 20 nM.

Los ensayos de retardo con DszR, se realizaron en un volumen final de 15 µl. La mezcla de reacción contenía: 20 fmoles de la sonda marcada (según el apartado IV.5.1 de esta sección), 100 ng de ADN de esperma de salmón, DTT 2 mM, 5 µg de BSA y cantidades crecientes de la proteína en tampón de unión (Tris-HCl 10 mM pH 8, NaCl 50 mM, EDTA 1 mM, glicerol 5 % (v/v).), la mezcla se incubó 20 minutos a temperatura ambiente, tras lo cual, se adicionaron 3 µl de tampón de carga 6X (azul de bromofenol 0,05% (p/v), xylene cyanol 0,05% (p/v), Tris-HCl 10 mM, EDTA 1 mM, glicerol 30% (pH 8)), para parar la reacción.

Para resolver las reacciones de unión ADN-proteína se realizó electroforesis en gel nativo de acrilamida de 1,5 mm al 5 % (p/v) en tampón TBE 0,5X (Tris-ácido bórico 45 mM, EDTA 1 mM (pH 8,3)) a 4°C. Los geles se secaron con un secador Hoefer™ Slab Gel Dryer CD 2000 (GE healthcare) y se expusieron a una pantalla sensible a radiaciones gamma (Storage Phosphor Screen, GE Healthcare). Las bandas se visualizaron con un escáner Typhoon 9410, pero no se realizó un análisis cuantitativo de las imágenes, ya que la finalidad del ensayo era comprobar si DszR podía unirse a ADN, como paso previo al ensayo de protección frente a DNasa I.

Para localizar *in vitro* las secuencias diana de DszR, se realizaron ensayos de protección frente a DNasa I, según lo descrito por Porrúa (Porrúa *et.al*, 2007), con algunas modificaciones. Las sondas se marcaron como se indica en IV.5.a. Las reacciones de unión siguiendo el protocolo descrito por Rivas-Marín (Rivas-Marín *et al.*, 2016), con algunos cambios en el tampón de unión. Se llevaron a cabo en 15 µl, con 20 fmoles de la sonda marcada, 100 ng de ADN de esperma de salmón, DTT 2 mM, 5 µg de BSA y cantidades crecientes de la proteína en tampón de unión ((Tris-HCl 10 mM pH 8, NaCl 50 mM, DTT 0,4 mM, glicerol 2 % (v/v), 2 mM CaCl₂, 5 mM MgCl₂). La mezcla se incubó 20 minutos a temperatura ambiente. Tras lo cual, se añadió 1 µl de una dilución en H₂O bidestilada (determinada empíricamente), de DNasa I no recombinante (Invitrogen), (diluciones 1/750 y 1/1000, aunque se optó finalmente por la dilución 1/1000) La mezcla se incubó durante 30 segundos y se paró la reacción añadiendo 4 µl de tampón de parada (acetato sódico 1,5 M pH 5,2, EDTA 130 mM, ADN esperma de salmón 1 mg/ml, glucógeno 2,4 mg/ml). Los tubos se transfirieron inmediatamente a hielo y se precipitaron las reacciones añadiendo 46 µl de etanol 96% (v/v) frío e incubándose al menos 12 horas a -20°C. Tras la precipitación, se centrifugaron a 20800 g y 4°C durante 30 minutos, se secaron los precipitados 5 minutos en termobloque a 45°C y 2 minutos a temperatura ambiente y se resuspendieron en 5 µl de tampón de carga (azul de bromofenol 0,125 % (p/v), xilencianol 0,125 % (p/v), EDTA 20 mM, formamida 95 % (v/v)).

Como marcador de tamaño se emplearon reacciones de secuenciación realizadas con el kit Sequenase 2.0 (USB), siguiendo las instrucciones del fabricante y usando 5 µg de ADN del plásmido como molde y 0,5 pmol del oligonucleótido específico (SepFpDszE1 para la cadena+). Tanto las muestras como las reacciones de secuenciación se hirvieron durante 5 minutos, se transfirieron a hielo y se cargaron en un gel de secuenciación de 0,2 mm TBE/acrilamida 6 % (p/v)-urea 8,3 M precorrido

previamente en TBE 1X. Las reacciones de secuenciación se corrieron en paralelo a las de *footprinting*. Los geles se secaron con un secador Hoefer™ Slab Gel Dryer CD 2000 (GE healthcare) y se expusieron a una pantalla sensible a radiaciones gamma (Storage Phosphor Screen, GE Healthcare). Las bandas se visualizaron con un escáner Typhoon 9410 (GE Healthcare).

RESULTADOS

I. Caracterización de los genes catabólicos *dszEABC*.

I.1. Aislamiento y caracterización del crecimiento del clon UPO112.

I.1.1. Búsqueda de un clon capaz de conferir a *E. coli* la capacidad de crecer en DBT como fuente de azufre.

En el trabajo de Tesis de la Dra. Laura Terrón, se construyó una metagenoteca, que contiene 6.5 Gb de ADN, (185000 clones) obtenido de muestras de pilas de refinería de una industria petroquímica (Terrón-González *et al.*, 2016). En dicha metagenoteca, la Dra. Terrón realizó una búsqueda de clones con actividad oxigenasa inicial (u oxigenasa de anillos aromáticos), uno de los cuales (UPO21), que sólo mostró actividad productora de índigo en presencia de arabinosa y salicilato, se secuenció completamente, encontrándose en su secuencia un operón de degradación de DBT (*dszABC*), cuyo gen *dszC* parecía ser el responsable de la producción de índigo. Los genes de biodesulfuración de DBT, venían precedidos de una FMN reductasa, que presumiblemente, sería el transportador de electrones a las oxigenasas, y que se denominó DszE. También se localizó en la secuencia de UPO21 el gen que codifica para el posible activador transcripcional de la ruta (DszR).

Puesto que el procedimiento de rastreo había consistido en un método identificativo indirecto tedioso y poco eficaz, con objeto de identificar más clones que pudieran albergar genes de degradación de DBT, en el presente trabajo de Tesis se decidió realizar una búsqueda diferente, utilizando un método selectivo para la capacidad de crecer con DBT como única fuente de azufre, y empleando para ello medio mínimo con dicha fuente de azufre.

Para poder realizar este rastreo, y puesto que la metagenoteca se mantiene en la estirpe EPI300-T1, en primer lugar, se transfirió por conjugación triparental a la estirpe MPO554 Nal^R, para maximizar la probabilidad de expresión de los genes presentes en los fósmidos, y aumentar, por tanto, la probabilidad de encontrar fósmidos de interés. La conjugación se realizó tal y como se describe en materiales y métodos (M&M). El parche de conjugación se diluyó 10 veces y se sembraron placas de medio mínimo M9 (descrito en M&M), con DBT como única fuente de azufre, con arabinosa para aumentar el nº de copias de los fósmidos, y salicilato para inducir la expresión heteróloga desde el promotor *P_{sal}*. Con objeto de seleccionar los clones que proporcionasen mejor crecimiento, las placas se incubaron a 30°C solo durante 48h, cuando se observó la aparición de microcolonias, 100 de las cuales se pusieron en placa de LB Cm12,5 µg/ml y Nal 15µg/ml. De ellas, se seleccionaron 20 candidatos que se inocularon en medio mínimo líquido con DBT como única fuente de azufre y en el mismo medio sin azufre

(pero añadiendo NN-dimetilformamida, solvente de DBT), y se incubaron a 30°C durante 48 h. Los 20 candidatos confirmaron su capacidad de crecer en DBT alcanzando una DO₆₀₀ en torno a 0,6 frente a 0,1 de los cultivos sin azufre, tal y como se muestra en la tabla 4.

| | DBT como fuente de S | sin azufre |
|--------------|----------------------|------------|
| candidato 1 | 0,66 | 0,1 |
| candidato 2 | 0,69 | 0,14 |
| candidato 3 | 0,59 | 0,11 |
| candidato 4 | 0,68 | 0,12 |
| candidato 5 | 0,84 | 0,13 |
| candidato 6 | 0,44 | 0,14 |
| candidato 7 | 0,44 | 0,09 |
| candidato 8 | 0,62 | 0,13 |
| candidato 9 | 0,55 | 0,12 |
| candidato 10 | 0,6 | 0,14 |
| candidato 11 | 0,55 | 0,11 |
| candidato 12 | 0,54 | 0,09 |
| candidato 13 | 0,67 | 0,12 |
| candidato 14 | 0,52 | 0,1 |
| candidato 15 | 0,89 | 0,13 |
| candidato 16 | 0,67 | 0,14 |
| candidato 17 | 0,78 | 0,15 |
| candidato 18 | 0,55 | 0,13 |
| candidato 19 | 0,65 | 0,14 |
| candidato 20 | 0,64 | 0,11 |

Tabla 4: DO₆₀₀ tras 48 horas de crecimiento, de los candidatos obtenidos en una búsqueda en una metagenoteca de pilas de una refinería empleando medio mínimo con DBT como única fuente de azufre.

Se extrajo ADN de los candidatos y se realizó un análisis de restricción con *Bam*HI, comparándose el patrón de bandas obtenido, entre ellos y con UPO21. Todos los candidatos presentaron el mismo patrón de restricción, el cual era diferente al de UPO21, por lo que se seleccionaron 2 de ellos para un análisis de restricción con *Hind*III y *Stu*I, que dieron el mismo patrón de bandas, lo que confirmó que se trataba de un único clon, efectivamente diferente a UPO21, y que se denominó UPO112.

I.1.2. Caracterización del crecimiento del clon UPO112.

En primer lugar, se procedió a caracterizar en detalle la capacidad de la estirpe MPO554 NaI^R con el fósido UPO112 de crecer en DBT como única fuente de azufre, y la dependencia o no, de este crecimiento de los inductores arabinosa y/o salicilato. Para ello se cultivó dicha estirpe en medio mínimo M9 (descrito en M&M), y se observó que *E. coli* con el fósido UPO112 puede crecer con DBT como única fuente de azufre mucho más rápido y alcanzando mayor DO, que con el clon ya descrito UPO21 (Martín-Cabello *et al.*, 2020); Además, el crecimiento es independiente de la adición de

arabinosa y salicilato, de hecho, la densidad óptica alcanzada con la adición de uno o los dos inductores fue menor (figura 3), lo que sugiere que los genes de degradación de DBT en UPO112 pueden expresarse por sí mismos a un nivel suficiente para permitir a *E. coli* un rápido crecimiento en DBT como única fuente de azufre.

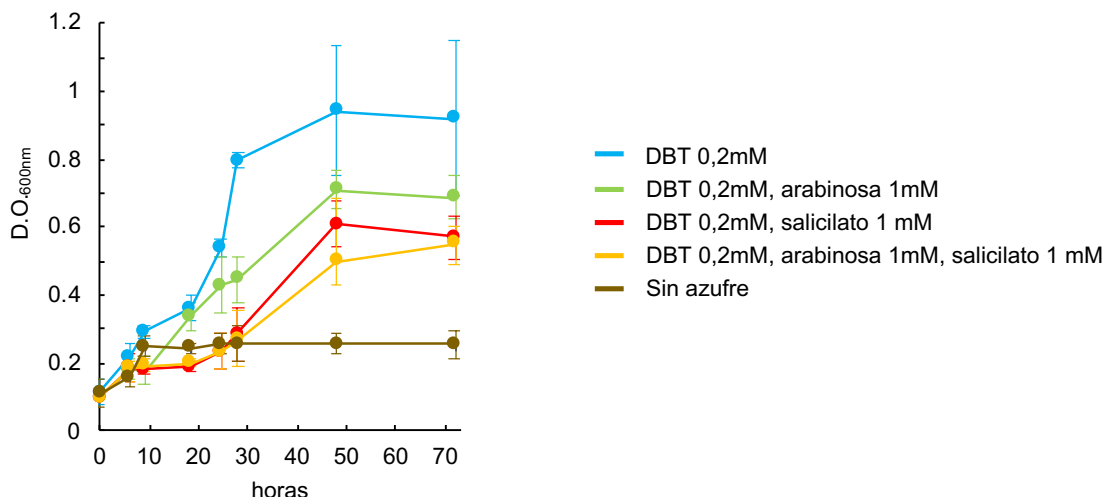


Figura 3: Curvas de crecimiento del clon UPO112 en medio mínimo con DBT 0,2 mM como fuente de azufre y con adición o no de inductores. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

Para ampliar el espectro de posibilidades de utilización de los genes presentes en UPO112, se chequeó también si *E. coli* con este fósido podía crecer con derivados metilados del DBT como 4,6-dimetil-dibenzotiofeno, presente también en crudos de petróleo y aún más recalcitrante que el DBT, y comparar la capacidad de crecimiento en dicho compuesto, con la capacidad de crecimiento en sulfato, su fuente de azufre convencional, así como en taurina, un sulfonato alifático que puede ser utilizado por *E. coli* como fuente de azufre, aunque menos preferencial que el sulfato, y con DBT. En la figura 4 podemos observar que, aunque la fuente de azufre en la que crece de forma más eficiente es sulfato, las curvas de crecimiento con sulfato y taurina son prácticamente similares, alcanzando en ambos casos una DO_{600} de 1 a las 24 horas, con tiempos de generación en torno a las 4 horas. Con DBT alcanza una DO_{600} aproximada de 0,8, al cabo de las 28 h, con un tiempo de generación de 8 horas aproximadamente, mientras que con 4,6-dimetil-dibenzotiofeno, a pesar de ser un compuesto muy recalcitrante, alcanza una DO_{600} de 0,56, al cabo de las 28h (tiempo de generación aproximado de 16 horas).

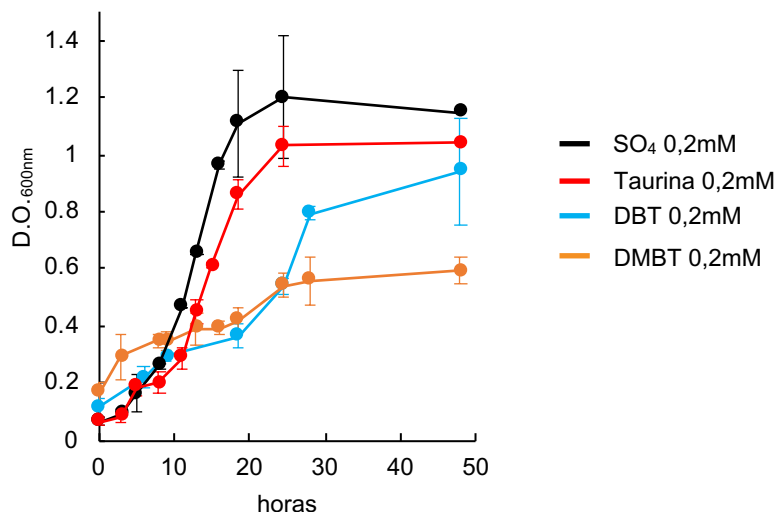


Figura 4: Curvas de crecimiento del clon UPO112 en medio mínimo con diferentes fuentes de azufre (0,2 mM de cada una de ellas). Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

Se comprobó también si UPO112 podía crecer más eficientemente aumentando la concentración de DBT o DMBT en el medio, de 0,2 mM a 1 mM, pero tal y como se muestra en la figura 5 no se observaron variaciones importantes, lo que indica que la limitación en el crecimiento a expensas de DBT o DMBT no es debido a falta de azufre.

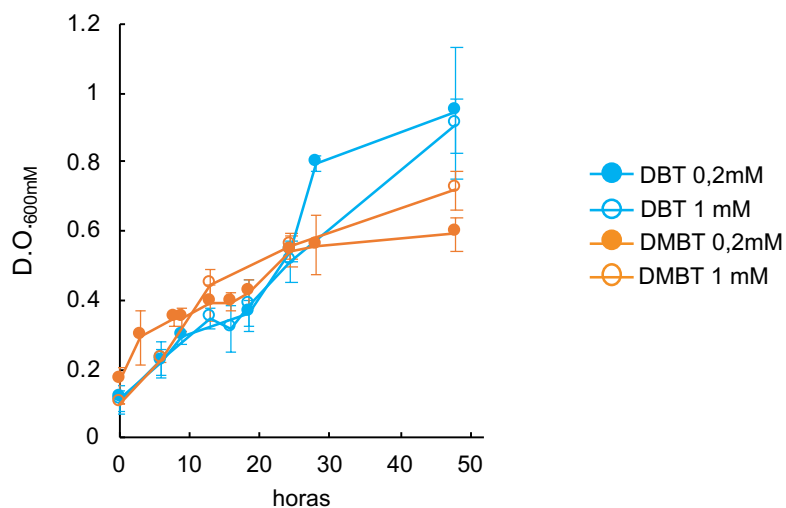


Figura 5: Curvas de crecimiento del clon UPO112 en medio mínimo con diferentes fuentes de azufre. Círculos cerrados 0,2 mM, y círculos abiertos 1 mM de cada una de ellas. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

Para determinar si el DBT se metaboliza por la vía 4S, (Piddington *et al.*, 1995), tal y como sucede en el clon UPO21, se ensayó la producción de 2-HBF, de UPO112

creciendo en medio mínimo con DBT como única fuente de azufre. La determinación se realizó mediante el método Gibbs, (Murarka *et al.*, 2019), tal y como se describe en materiales y métodos, obteniéndose un resultado de 0,18 mM de 2-HBF producido a partir de 0,2 mM de DBT, al cabo de las 48 horas, lo cual muestra una eficiencia mayor con respecto a UPO21 (0,15 mM producido tras 7 días; Martín-Cabello *et al.*, 2020), y correlaciona con su mejor crecimiento.

Dado que UPO21 se había aislado en una búsqueda de clones con actividad oxigenasa hidroxilante de anillos aromáticos (Terrón, 2014), gracias a su capacidad de producir índigo, el cual presenta color azul, a partir del triptófano del medio LB, se comprobó si UPO112 también tenía esta capacidad, y se observó que UPO112 no produce índigo en medio LB, ni con arabinosa y salicilato presentes, ni sin ellos, tal y como se observa en la figura 6.

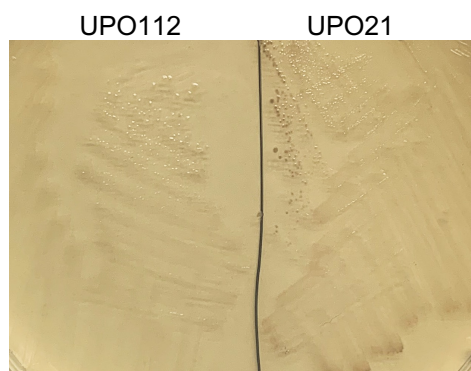


Figura 6: Producción de índigo por parte de los clones UPO21 y UPO112, en medio LB con arabinosa y salicilato.

Debido a estos resultados, diferentes a los obtenidos con el clon previamente caracterizado, se decidió estudiar con mayor profundidad el clon UPO112, y para ello se secuenció completamente.

I.2. Análisis de secuencias de UPO112.

I.2.1. Asignación de *orfs*.

El clon UPO112 se secuenció completamente, obteniéndose un inserto de 36286 pb, en el cual hay 36 *orf* completas. Este inserto es ligeramente mayor que el de UPO21, que contenía 34226 pb, e incluía 31 *orf* completas.

En la tabla 5 se muestran los resultados de comparar las proteínas codificadas por cada *orf* de UPO112 con las bases de datos: “non-redundant protein sequences” (nr), y UniProtKB/Swiss-Prot (swissprot). Con la segunda base de datos, se obtienen algunos resultados cuyos e-value son demasiado altos como para ser significativos, y otros en

los que la mayor similitud se da con organismos superiores, con lo que, por lo general, los resultados que se han tenido en cuenta son los obtenidos con la base de datos nr.

| | Database: non-redundant protein sequences (nr), id | Database: UniProtKB/Swiss-Prot (swissprot), id |
|--------------|---|---|
| <i>orf1*</i> | 4a-hydroxytetrahydrobiopterin dehydratase [Candidatus Macondimonas diazotrophica], 77,67% | Full=Putative pterin-4- α -carbinolamine dehydratase; Short=PHS; AltName: Full=4- α -hydroxy-tetrahydropterin dehydratase; AltName: Full=Pterin carbinolamine dehydratase; Short=PCD [Nitrosomonas eutropha C91], 60,52% |
| <i>orf2</i> | oligoribonuclease [Candidatus Macondimonas diazotrophica], 84,82% | Full=Oligoribonuclease [Pseudomonas stutzeri A1501], 71,10% |
| <i>orf3</i> | DMT family transporter [Candidatus Macondimonas diazotrophica], 70,59% | Pseudopaline exporter Cntl [Pseudomonas aeruginosa PAO1], 32,51% |
| <i>orf4</i> | phosphoenolpyruvate synthase [Candidatus Macondimonas diazotrophica], 91,27% | Full=Phosphoenolpyruvate synthase; Short=PEP synthase; AltName: Full=Pyruvate, water dikinase [Pseudomonas aeruginosa UCBPP-PA14], 73,12% |
| <i>orf5</i> | kinase/pyrophosphorylase [Candidatus Macondimonas diazotrophica], 91,61% | Full=Putative phosphoenolpyruvate synthase regulatory protein; Short=PEP synthase regulatory protein; Short=PSRP; AltName: Full=Pyruvate, water dikinase regulatory protein [Pseudomonas aeruginosa LESB58], 53,11% |
| <i>orf6</i> | DUF4345 family protein [Candidatus Macondimonas diazotrophica], 62,80% | Full=Protein piccolo; AltName: Full=Aczonin; AltName: Full=Brain-derived HLMN protein; AltName: Full=Multidomain presynaptic cytomatrix protein [Mus musculus], 45,45% |
| <i>orf7</i> | ferritin [Candidatus Macondimonas diazotrophica], 87,25% | Full=Ferric enterobactin transport protein FepE [Escherichia coli K-12], 46,43% |
| <i>orf8</i> | DUF1249 domain-containing protein [Candidatus Macondimonas diazotrophica], 65,33% | Full=Uncharacterized protein YqiB [Escherichia coli K-12], 28,15% |
| <i>orf9</i> | MaoC family dehydratase [Candidatus Macondimonas diazotrophica], 92,16% | Full=Probable enoyl-CoA hydratase 1; AltName: Full=N-related protein; AltName: Full=Nodulation protein [Mycobacterium tuberculosis CDC1551], 46,98% |
| <i>orf10</i> | uroporphyrinogen-III synthase [Candidatus Macondimonas diazotrophica], 76,57% | Full=Uncharacterized protein YjjA [Bacillus subtilis subsp. subtilis str. 168], 30,45% |
| <i>orf11</i> | cardiolipin synthase [Candidatus Macondimonas diazotrophica], 78,83% | Full=Major cardiolipin synthase CIsA; Short=CL synthase 2 [Bacillus subtilis subsp. subtilis str. 168], 32,40% |
| <i>orf12</i> | NAD(P)-dependent alcohol dehydrogenase [Candidatus Macondimonas diazotrophica], 74,85% | Full=Zinc-type alcohol dehydrogenase-like protein C1773.06c [Schizosaccharomyces pombe 972h-], 39,30% |
| <i>orf13</i> | FAD-binding oxidoreductase [Candidatus Macondimonas diazotrophica], 70,29% | Full=D-2-hydroxyglutarate dehydrogenase, mitochondrial; Short=AtD-2HGDH; Flags: Precursor [Arabidopsis thaliana], 33,69% |
| <i>orf14</i> | sulfate permease [Candidatus Macondimonas diazotrophica], 75,13% | Full=Sulfate transporter 4.1, chloroplastic; AltName: Full=AST82; Flags: Precursor [Arabidopsis thaliana], 32,57% |
| <i>orf15</i> | MBL fold metallo-hydrolase [Candidatus Macondimonas diazotrophica], 89,91% | Full=Persulfide dioxygenase ETHE1, mitochondrial; AltName: Full=Ethylmalonic encephalopathy protein 1 homolog; AltName: Full=Hepatoma subtracted clone one protein; AltName: Full=Sulfur dioxygenase ETHE1; Flags: Precursor [Mus musculus], 47,37% |

| | | |
|---------------------|---|--|
| <i>orf16</i> | hypothetical protein [Candidatus Macondimonas diazotrophica], 58,83% | No existen similitudes significativas. |
| <i>orf17</i> | methyltransferase domain-containing protein [Candidatus Macondimonas diazotrophica], 82,61% | Full=Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE; AltName: Full=2-methoxy-6-polyprenyl-1,4-benzoquinol methylase; AltName: Full=Demethylmenaquinone methyltransferase [Geobacter daltonii FRC-32], 35,03% |
| <i>orf18</i> | phospholipase [Candidatus Macondimonas diazotrophica], 71,08% | Full=Putative ATP-dependent RNA helicase PB1A10.06c [Schizosaccharomyces pombe 972h-], 29,17% |
| <i>orf19</i> | acetylglutamate kinase [Candidatus Macondimonas diazotrophica], 90,46% | Full=Acetylglutamate kinase; AltName: Full=N-acetyl-L-glutamate 5-phosphotransferase; AltName: Full=NAG kinase; Short=NAGK [Alkalilimnicola ehrlichii MLHE-1], 75,17% |
| <i>orf20</i> | DUF4124 domain-containing protein [Candidatus Macondimonas diazotrophica], 55,50% | Full=Formin BNR1 [Candida albicans SC5314], 31,71% |
| <i>orf21</i> | orotate phosphoribosyltransferase [Candidatus Macondimonas diazotrophica], 79,34% | Full=Orotate phosphoribosyltransferase; Short=OPRT; Short=OPRTase [Marinomonas sp. MWYL1], 60,28% |
| <i>orf22 - dszH</i> | alpha/beta hydrolase [Propionivibrio dicarboxylicus], 54,88% | No existen similitudes significativas, |
| <i>orf23 - dszR</i> | DszR [uncultured bacterium], (UPO21) 72,41% | Full=Sigma54-dependent transcriptional regulator SfnR [Pseudomonas fluorescens Pf0-1], 55,69% |
| <i>orf24</i> | hydroxyacylglutathione hydrolase [uncultured bacterium] (UPO21) , 65,92% | Full=Methylthioribose kinase; Short=MTR kinase [Yersinia pestis Pestoides F], 29,41% |
| <i>orf25</i> | OmpW family protein [uncultured bacterium] (UPO21), 69,14% | No existen similitudes significativas. |
| <i>orf26</i> | DUF2141 domain-containing protein [uncultured bacterium] (UPO21) , 57,64% | Full=Protein sidekick-2; Flags: Precursor [Gallus gallus], 34,25% |
| <i>orf27</i> | hypothetical protein [uncultured bacterium] (UPO21), 69,80% | Full=Protein sidekick-2; Flags: Precursor [Gallus gallus], 38,46% |
| <i>orf28 - dszC</i> | DszC [uncultured bacterium], (UPO21) 90,56% | Full=Dibenzothiophene desulfurization enzyme C; AltName: Full=DBT sulfur dioxygenase [Rhodococcus sp. IGTS8], 78,42% |
| <i>orf29 - dszB</i> | ABC transporter substrate-binding protein (Comamonadaceae bacterium), 90, 11% | Full=2'-hydroxybiphenyl-2-sulfinate desulfinase; AltName: Full=Dibenzothiophene desulfurization enzyme B [Rhodococcus sp. IGTS8], 71,78% |
| <i>orf30 - dszA</i> | DszA [uncultured bacterium], (UPO21) 89,87% | Full=Dibenzothiophene desulfurization enzyme A [Rhodococcus sp. IGTS8], 81,15% |
| <i>orf31 - dszE</i> | NADPH-dependent FMN reductase [Comamonadaceae bacterium], 93,75% | Full=FMN reductase (NADPH); AltName: Full=FMN reductase [Pseudomonas putida KT2440], 40,98% |
| <i>orf32*</i> | IS3 family transposase [Xanthomonas arboricola pv. juglandis], 78,72% | Full=Insertion element IS407 uncharacterized 31.7 kDa protein; AltName: Full=ORF1 [Burkholderia multivorans ATCC 17616], 43,48% |
| <i>orf33</i> | transposase [Polycyclovorans sp.], 91,09%, | Full=Insertion element IS407 uncharacterized 31.7 kDa protein; AltName: Full=ORF1 [Burkholderia multivorans ATCC 17616], 27,61% |
| <i>orf34*</i> | transposase [Arenimonas composti], 91,94% | Full=Insertion element IS407 uncharacterized 31.7 kDa protein; AltName: Full=ORF1 [Burkholderia multivorans ATCC 17616], 29,09% |

| | | |
|---------------|--|--|
| <i>orf35*</i> | hypothetical protein PS903_05943 [Pseudomonas fluorescens], 75,86% | No existen similitudes significativas. |
| <i>orf36</i> | NAD-dependent DNA ligase [bacterium], 77,29% | Full=DNA ligase; AltName: Full=Polydeoxyribonucleotide synthase [NAD(+)] [uncultured Termite group 1 bacterium phylotype Rs-D17], 34,91% |
| <i>orf37</i> | hypothetical protein [Nitrosomonas mobilis], 43,09% | No significant similarity found. |
| <i>orf38*</i> | phage integrase N-terminal SAM-like domain-containing protein [Gammaproteobacteria bacterium] 70,97% | Full=Integrase/recombinase; AltName: Full=E2 protein [Escherichia coli], 40,54% |
| <i>orf39</i> | hypothetical protein [Immundisolibacter cernigliae], 73,79% | No existen similitudes significativas. |
| <i>orf40</i> | hypothetical protein [Candidatus Macondimonas diazotrophica], 62,02% | Full=Thiosulfate sulfurtransferase GlpE [Edwardsiella ictaluri 93-146], 42,57% |
| <i>orf41</i> | hypothetical protein [Immundisolibacter cernigliae], 79,35% | No existen similitudes significativas. |
| <i>orf42*</i> | hypothetical protein [Candidatus Macondimonas diazotrophica], 77,91% | Full=GTP 3',8-cyclase; AltName: Full=Molybdenum cofactor biosynthesis protein A [Alkaliphilus oremlandii OhILAs], 47,62% |

Tabla 5: resultados de comparar las proteínas codificadas por cada orf de UPO112 con las bases de datos: “non-redundant protein sequences” (nr) (blastp), y UniProtKB/Swiss-Prot (swissprot). (* indica proteína incompleta)

Podemos observar que las proteínas a las que más se parecen las codificadas por las *orf1* (incompleta) hasta la *orf21* pertenecen a *Candidatus Macondimonas diazotrophica* (Karthikeyan *et al.*, 2019).

La proteína codificada por *orf22*, a la que más se parece es a una alfa/beta hidrolasa de *Propionivibrio dicarboxylicus*, pero sólo tiene una identidad con ella del 54,88% (con una cobertura del 75% y e value del 9e-50), esta búsqueda, se realizó con blastp, ya que la opción quickblastp utilizada por defecto para la realización de análisis de secuencia en este trabajo de Tesis, no arrojó ningún resultado, puesto que dicha opción, incluye un preprocesamiento de la base de datos nr, que implica mostrar las secuencias con mayor similitud. Puesto que el gen *orf 22* solapa con *dszR*, se comparó la proteína codificada por *orf22* con la proteína codificada por el gen aguas arriba de *dszR* en UPO21 (*gen11*), observándose que presentan cierta identidad (37,63% con una cobertura del 92%). Debido al solapamiento con *dszR* y a su implicación en la expresión del operón *dszEABC* (ver más adelante), este gen se ha designado *dszH*. *orf23* codifica para una proteína que a la que más se parece es a DszR de UPO21, que codifica para un activador transcripcional dependiente de σ^N (inicialmente descrito en Terrón, 2014 y completamente caracterizado posteriormente en esta Tesis) (con una identidad del 72,4% y una cobertura del 100%).

Las *orfs* 24-25-26 y 27 codifican para proteínas que a las que más se parecen son proteínas del fósido UPO21, (con identidades entre el 57 y el 69% y coberturas entre el 82-100%).

En cuanto al grupo de *orfs*28-29-30-31, podría constituir el operón *dszEABC*_{UPO112}, similar al de UPO21, descrito en la Tesis de la Dra. Terrón. Las proteínas codificadas por dichos genes en UPO21, forman parte de la ruta de desulfuración de dibenzotiofeno, vía 4S. Prestando atención a cada una de estas *orfs*; se observa que en el caso de *orf* 28, la proteína con la que presenta mayor similitud es DszC_{UPO21} (Dibenzotiofeno monooxigenasa) con una identidad aproximada del 91 % (cobertura 100%). A continuación, *orf*29 codifica para una proteína cuya similitud con DszB_{UPO21} (2-hidroxibifenilo-2-sulfinato desulfinasa) es del \approx 82%, con una cobertura del 100%; sin embargo, esta no es la opción con mayor puntuación en la base de datos puesto que *orf*29, presenta una similitud del \approx 90%, con una proteína anotada como la proteína de unión al sustrato de un transportador tipo ABC de *Comamonadaceae bacterium* (también con cobertura del 100%). Si tenemos en cuenta en este caso los resultados de la base de datos Uniprot/Swissprot, podemos observar que la proteína codificada por *orf*29, a la que más se parece es a una 2-hidroxibifenilo-2-sulfinato desulfinasa, (de *Rhodococcus sp. IGTS8*). La siguiente *orf*, (*orf*30), codifica para una proteína cuya mayor similitud (del \approx 90%,) es con DszA_{UPO21} (Dibenzotiofeno-5,5-dióxido monooxigenasa). Por último, *orf*31, presenta una similitud de casi el 94% (y cobertura del 100%) con una FMN reductasa dependiente de NADPH, de *Comamonadaceae bacterium*. Pero, además, con la misma cobertura, presenta similitud del \approx 71% con DszE_{UPO21} (FMN reductasa).

La proteína codificada por *orf*33, a la que más se parece es a una transposasa, mientras que las *orf*32 y 34 parecen ser restos de ellas.

Las *orf*35 (incompleta) y 37 son proteínas hipotéticas.

*orf*36 codifica para una proteína cuya mayor similitud es con una ligasa dependiente de NAD.

*orf*38, parece ser un fragmento de una integrasa.

Las proteínas codificadas por *orf*40 y 42 (incompleta), presentan la mayor similitud con proteínas que pertenecen a *Candidatus Macondimonas diazotrophica*. En cuanto a *orf*39 y 41, la tercera opción con mayor puntuación de similitud, con coberturas iguales o superiores a la opción con mayor puntuación, y con alta identidad (\approx 73-70%), respectivamente, se da también con proteínas de *Candidatus Macondimonas diazotrophica*.

Cabe destacar que algunas proteínas, como *orf14* y *orf40* podrían estar relacionadas de alguna manera con transporte o metabolismo de azufre.

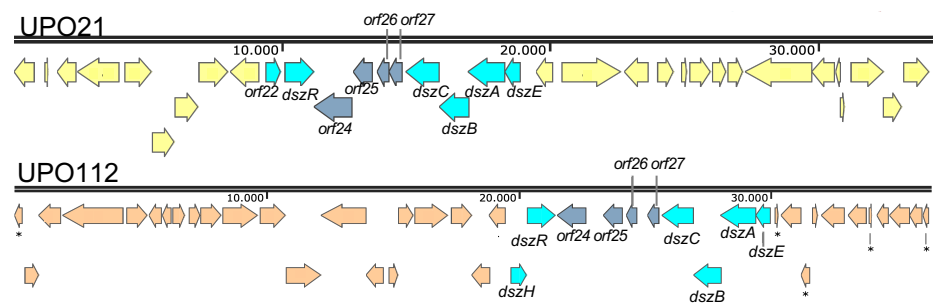


Figura 7: Esquema de la organización génica de los fósmidos UPO21 y UPO112. En celeste, genes *dsz*. En azul, genes similares en ambos fósmidos no relacionados con degradación de DBT. En otros colores, genes específicos de cada fósrido.

1.2.2.Comparativa con UPO21 y *Comamonadaceae bacterium*.
Si comparamos en detalle las secuencias de los genes de UPO112 que codifican para proteínas que presentan similitud con proteínas de UPO21 (GenBank: MK510930), podemos observar que en general, tal y como se muestra en las tablas 6, 7 y 8, presentan una alta identidad de secuencia y una estructura génica similar, con ciertas consideraciones para tener en cuenta.

| | nº nt UPO21 | nº nt UPO112 | nº aa UPO21 | nº aa UPO112 |
|-------------------|-------------|--------------|-------------|--------------|
| <i>gen11/dszH</i> | 576 | 651 | 191 | 216 |
| <i>dszR</i> | 1128 | 1134 | 375 | 377 |
| <i>orf24</i> | 1407 | 1161 | 468 | 386 |
| <i>orf25</i> | 744 | 747 | 247 | 248 |
| <i>orf26</i> | 435 | 435 | 144 | 144 |
| <i>orf27</i> | 477 | 477 | 158 | 158 |
| <i>dszC</i> | 1242 | 1239 | 413 | 412 |
| <i>dszB</i> | 1116 | 1116 | 371 | 371 |
| <i>dszA</i> | 1395 | 1395 | 464 | 464 |
| <i>dszE</i> | 579 | 579 | 192 | 192 |

Tabla 6: Comparativa tamaño genes y proteínas UPO21 vs UPO112

| | Región intergénica UPO21 | Región intergénica UPO112 | Región intergénica <i>Comamonadaceae bacterium</i> |
|----------------------|--------------------------|---------------------------|--|
| <i>gen 11/dszH-R</i> | 123 | Solapan 4 pb | |
| <i>dszR-orf24</i> | solapan33 | 22 | |
| 24-25 | 31 | 678 | |
| 25-26 | 159 | 162 | |
| 26-27 | 39 | 399 | |
| <i>orf27-dszC</i> | 120 | 123 | |
| <i>dszC-B</i> | solapan 4 | solapan 4 | solapan 3 |
| <i>dszB-A</i> | solapan 37 pb | solapan 37 pb | solapan 36 pb |
| <i>dszA-E</i> | 4 pb | 5 pb | 6 pb |

Tabla 7: Comparativa regiones intergénicas UPO21 vs UPO112 vs *Comamonadaceae bacterium*.

| | id (% nt) | cobertura (%) | E value | id (% aa) | cobertura (%) | E value |
|-------------------|-----------|---------------|---------|-----------|---------------|---------|
| <i>gen11/dszH</i> | 81,25 | 5 | 9e-05 | 37,63 | 92 | 2e-32 |
| <i>dszR</i> | 77,88 | 89 | 2e-179 | 72,41 | 100 | 0,0 |
| <i>orf24</i> | 71,63 | 79 | 3e-163 | 68,39 | 82 | 0,0 |
| <i>orf25</i> | 70,00 | 72 | 1e-68 | 69,14 | 98 | 3e-127 |
| <i>orf26</i> | 71,53 | 64 | 1e-37 | 57,64 | 100 | 3e-65 |
| <i>orf27</i> | 70,88 | 92 | 2e-54 | 69,80 | 94 | 2e-76 |
| <i>dszC</i> | 85,82 | 95 | 0,0 | 90,58 | 100 | 0,0 |
| <i>dszB</i> | 80,42 | 97 | 0,0 | 81,99 | 100 | 0,0 |
| <i>dszA</i> | 86,62 | 98 | 0,0 | 89,89 | 100 | 0,0 |
| <i>dszE</i> | 74,96 | 97 | 1e-104 | 71,50 | 100 | 1e-89 |

Tabla 8: Resultados de comparar *orfs* de UPO21 vs *orfs* UPO112 (blastn/blastp).

En el caso de los genes estructurales de degradación de DBT *dszA*, *dszB* y *dszC*, los de UPO112 presentan identidades superiores al 80% con los genes de UPO21 (con cobertura > 95%), mientras que la identidad y cobertura de *dszE*, son menores, como se detallará más adelante. Los tamaños de los cuatro genes, así como la organización génica de esta zona, en un bloque compacto, es prácticamente idéntica entre ambos clones, incluyendo el solapamiento existente entre los genes, que sugiere un acoplamiento traduccional (Tabla 6).

Puesto que *DszB* y *DszE* de UPO112, mostraron alta similitud con proteínas de *Comamonadaceae bacterium*, se realizó también un análisis comparativo con el genoma de dicha bacteria (JAECKR010000006.1), localizando en él un posible operón *dszEABC*, cuyos genes presentan una identidad mayor del 92%, con los de UPO112, con e-value de 0. La organización génica de *dszEABC* en este genoma es similar a la de UPO21 y UPO112, (tablas 7 y 9), siendo un elemento a destacar el hecho de que el operón se encuentre precedido igualmente de una reductasa.

| | alineamiento | id (% aa) | E value | Posiciones en genoma de <i>Comamonadaceae bacterium</i> |
|-------------|--------------|-----------|---------|---|
| <i>dszR</i> | 0 | | | No existen similitudes |
| <i>dszC</i> | gen completo | 93% | 0 | 50005 a 51243 |
| <i>dszB</i> | gen completo | 92% | 0 | 48884 a 50008 |
| <i>dszA</i> | gen completo | 93% | 0 | 47526 a 48920 |
| <i>dszE</i> | gen completo | 94% | 0 | 46942 a 47520 |

Tabla 9: resultados de comparar *orfs* de UPO112 vs *Comamonadaceae bacterium* (blastn).

Prestando especial atención a *dszE*, cabe destacar que *dszE*_{UPO21} y *dszE*_{UPO112} alinean a partir de su nucleótido 17, aunque la región promotora de ambos genes es similar, tal y como se abordará en el capítulo II. Si realizamos un alineamiento múltiple de las 3 proteínas con clustalomega (figura 8), podemos observar que la similitud en su secuencia comienza en la metionina inicial de UPO112, y la de UPO21, que corresponden con el residuo 25 de la proteína NADPH-dependent FMN reductase de *Comamonadaceae bacterium* (MBH2042284.1), a partir de ese aminoácido *DszE*_{UPO112}

y la proteína de *Comamonadaceae bacterium* son prácticamente idénticas en toda su longitud. Debemos tener en cuenta que al introducir la reductasa de *Comamonadaceae bacterium* en la base de datos “non-redundant protein sequences” (nr), las proteínas a las que más se parece (tras DszE_{UPO21}, que aparece en el primer lugar), son alcanosulfonatomonooxigenas, pero no se parece a ellas en su extremo amino, lo cual concuerda con que dicho extremo también sea diferente de los de DszE_{UPO112} y DszE_{UPO21}. En cuanto a la similitud entre DszE_{UPO112} y DszE_{UPO21}, observamos que, a pesar de que alinean desde la metionina inicial, tras ésta, hay una zona de aminoácidos no conservados o con sustituciones semiconservadas, y que, a partir del 7º residuo, hay una región altamente conservada. (Esto concuerda con que *dszE_{UPO112}* y *dszE_{UPO21}* alineen desde el nucleótido 17, como se ha comentado). Podemos destacar que la región central de las proteínas está altamente conservada en las tres secuencias. La similitud de DszE_{UPO21} con las otras 2 proteínas es un poco menor a partir de su residuo 147; y aunque el extremo carboxilo no es totalmente idéntico en las tres, si está altamente conservado.

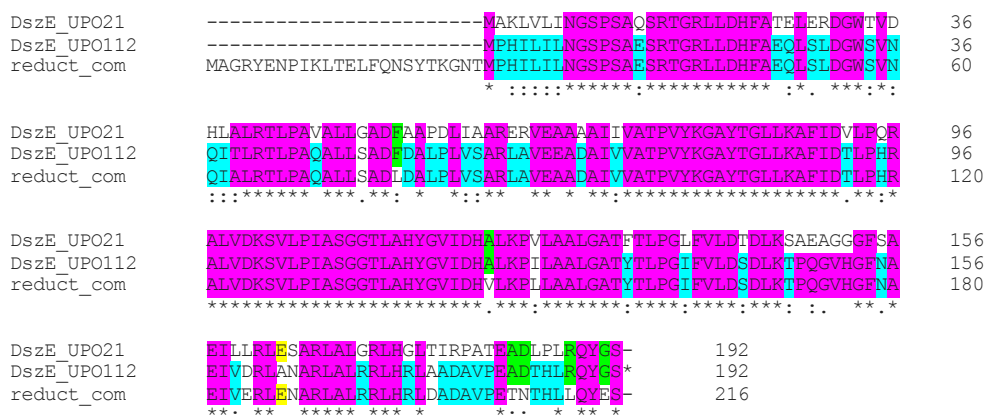


Figura 8: Alineamiento múltiple de DszE_{UPO21}, DszE_{UPO112} y NADPH-dependent FMN reductase de *Comamonadaceae bacterium* (MBH2042284.1), realizado con clustalomega. Se resaltan en color fucsia los residuos conservados en las 3 secuencias, en color azul los conservados entre DszE_{UPO112} y NADPH-dependent FMN reductase *Comamonadaceae bacterium*, en verde los conservados entre DszE_{UPO112} y DszE_{UPO21} y en amarillo los conservados entre DszE_{UPO21} y NADPH-dependent FMN reductase de *Comamonadaceae bacterium*.

DszE de UPO21, está definida como la FMN reductasa (NADH dependiente), que suministra los electrones a las monooxigenasas de la ruta de degradación de DBT (Terrón, 2014, Martín-Cabello *et al.*, 2020). Es importante destacar, que la región intergénica entre *dszE* y *dszA* tanto en UPO21, como en UPO112 y *Comamonadaceae bacterium* es muy pequeña (4, 5 y 6 nucleótidos respectivamente) (tabla 7), a diferencia de lo que sucede en otras rutas de degradación de DBT, como es el caso de

Rhodococcus sp. IGTS8, donde la reductasa de la ruta está codificada por *dszD*, que se encuentra alejado de los genes *dszABC* (Gray *et al.*, 1996). Si introducimos DszE (ya sea de UPO21 o de UPO112) en la base de datos UniProtKB/Swiss-Prot, obtenemos en ambos casos el mismo resultado, pudiendo comprobar que las proteínas bien caracterizadas a las que más se parece DszE, son reductasas dependientes de NADPH, pertenecientes a la familia SsuE. De hecho, al realizar un árbol filogenético con proteínas representativas de cada grupo con los que presentan similitud, e incluyendo en dicho análisis proteínas implicadas en la degradación de DBT, (figura 9), observamos que DszE_{UPO112} y la FMN reductasa de *Comamonadaceae bacterium* aparecen en 2 ramas que divergen muy poco entre sí, y muy cercanas a DszE_{UPO21}, y que ambas se encuentran alejadas de las FMN reductasas participes en la degradación de DBT, siendo más cercanas a las FMN reductasas asociadas a las sulfonato monooxigenasas alifáticas, y en particular las reductasas SsuE, que catalizan la desulfonación de diversos alcanosulfonatos. SsuE se ha caracterizado como una FMN reductasa dependiente de NADPH, necesaria para la utilización de dimetilsulfóxido en *P. putida* DS1, y cuya función es comparable a la de DszD (Endoh *et al.*, 2003b).

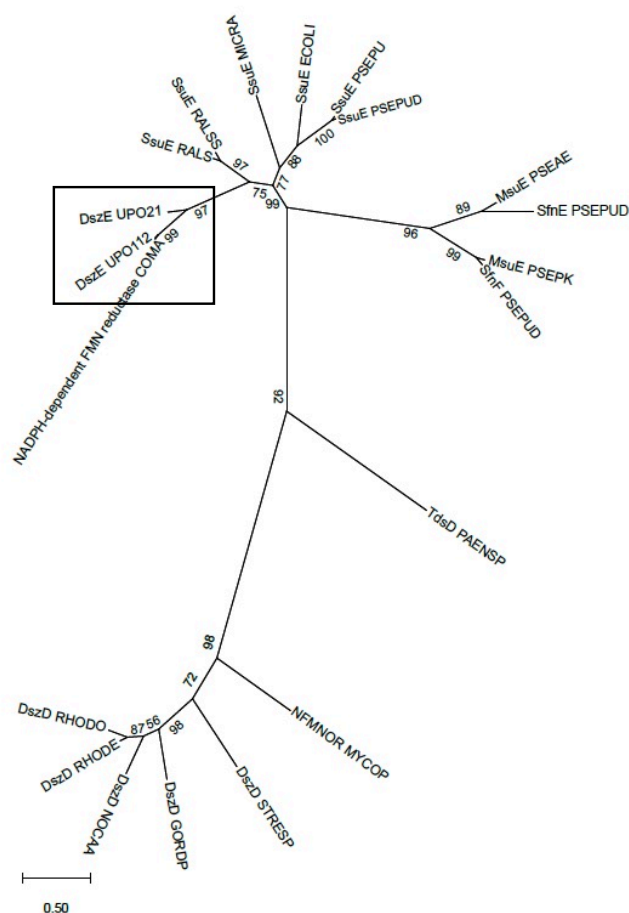


Figura 9: Árbol filogenético de DszE_{UPO112} y DszE_{UPO21} (se destacan con un recuadro estas 2 proteínas), junto con NADPH-dependent FMN reductase de *Comamonadaceae bacterium*

(MBH2042284.1) y las siguientes oxidorreductasas: SsuE_RALS, FMN reductasa (dependiente de NADPH) de *Ralstonia* (WP_045787849.1). SsuE_MICRA, FMN reductasa (dependiente de NADPH) de *Microcystis aeruginosa* (WP_012265981.1). SsuE_PSEPU, FMN reductasa (dependiente de NADPH) de *Pseudomonas putida* (O85762.2). SsuE_ECOLI, FMN reductasa (dependiente de NADPH) de *Escherichia coli* K-12 (P80644.2). MsuE_PSEAE, FMN reductasa (dependiente de NADPH) de *Pseudomonas aeruginosa* PAO1 (O31038.1). MsuE_PSEPK, FMN reductasa (dependiente de NADPH) de *Pseudomonas putida* KT2440 (Q88J85.1). SsuE_RALSS, FMN reductasa (dependiente de NADPH) de *Ralstonia solanacearum* (WP_039549499.1). DszD_STRESP, FMN reductasa (dependiente de NADH) DszD de *Streptomyces* sp. NBRC110027 (GAO11166.1). DszD_RHODE, DszD de *Rhodococcus erythropolis* (ABV44406.1). DszD_RHODO, FMN reductasa (dependiente de NADH) DszD de *Rhodococcus opacus* B4 (BAH52549.1). DszD_GORDP, FMN reductasa (dependiente de NADH) DszD de *Gordonia polyisoprenivorans* (GAB24076.1). DszD_NOCAA, FMN reductasa (dependiente de NADH) DszD de *Nocardia asteroides* NBRC 15531 (GAD84696.1). TdsD_PAENSP, oxidoreductasa (dependiente de NAD(P)H) de *Paenibacillus* sp. A11-2 (BAB13707.1). NFMNOR_MYCOP, FMN reductasa (dependiente de NADH) de *Mycobacterium phlei* (BAD83685.1). SfnE_PSEPUD, FMN reductasa (dependiente de NADH) de *Pseudomonas putida* DS1 (Q845S9.1). SsuE_PSEPUD, FMN reductasa (dependiente de NADPH) de *Pseudomonas putida* DS1 (BAC00971.1). SfnF_PSEPUD, FMN reductasa (dependiente de NADH) de *Pseudomonas putida* DS1 (BAD51729.1).

En cuanto a las *orfs* 24, 25, 26 y 27 de UPO112, que separan los genes *dsz* estructurales del potencial regulador *dszR*, presentan identidades en torno al 70%, (con cobertura entre el 64-92%) con las *orfs* correspondientes de UPO21. La organización génica de esta región es bastante parecida en ambos clones, aunque las distancias intergénicas por lo general son mayores en UPO112. Sin embargo, estas *orfs* no están presentes en el genoma de *Comamonadaceae bacterium*.

Prestando atención a *dszR*_{UPO112}, podemos observar que alinea con *dszR*_{UPO21}, en un 89% de su secuencia (desde el nucleótido 109 de UPO21 y desde el nucleótido 115 de UPO112 hasta 4pb antes del codón de STOP en ambos genes) con una identidad \approx 78%, (e value de $2e-179$) (tabla 8) y no se encuentran similitudes significativas entre *dszR* y el genoma de *Comamonadaceae bacterium*.

Al comparar DszR de UPO21 y de UPO112 con la base de datos Uniprot/Swissprot, observamos que muestran gran similitud con SfnR de *Pseudomonas fluorescens* (Q3K999.1) (identidad del 57,99%, con una cobertura del 88%, e value $1e-130$, en el caso de UPO21 e identidad del 55,69%, con cobertura del 89% y e value $2e-128$ en el caso de UPO112) (Martín-Cabello *et al.* 2020, tabla 5). Al realizar un alineamiento múltiple de las 3 proteínas con clustalomega, podemos observar que las mayores

diferencias entre ellas se encuentran en su extremo amino, y que la parte central está muy conservada entre las 3 (figura 10). SfnR, es un activador transcripcional implicado en la utilización de sulfuro de dimetilo como fuente de azufre, bajo el control de CysB, regulador transcripcional tipo LysR (Endoh, 2003) que también carece del dominio amino-terminal. Tanto SfnR como DszR podrían clasificarse dentro del grupo IV de activadores transcripcionales dependientes de σ^N , según la clasificación de Bush & Dixon, 2012.

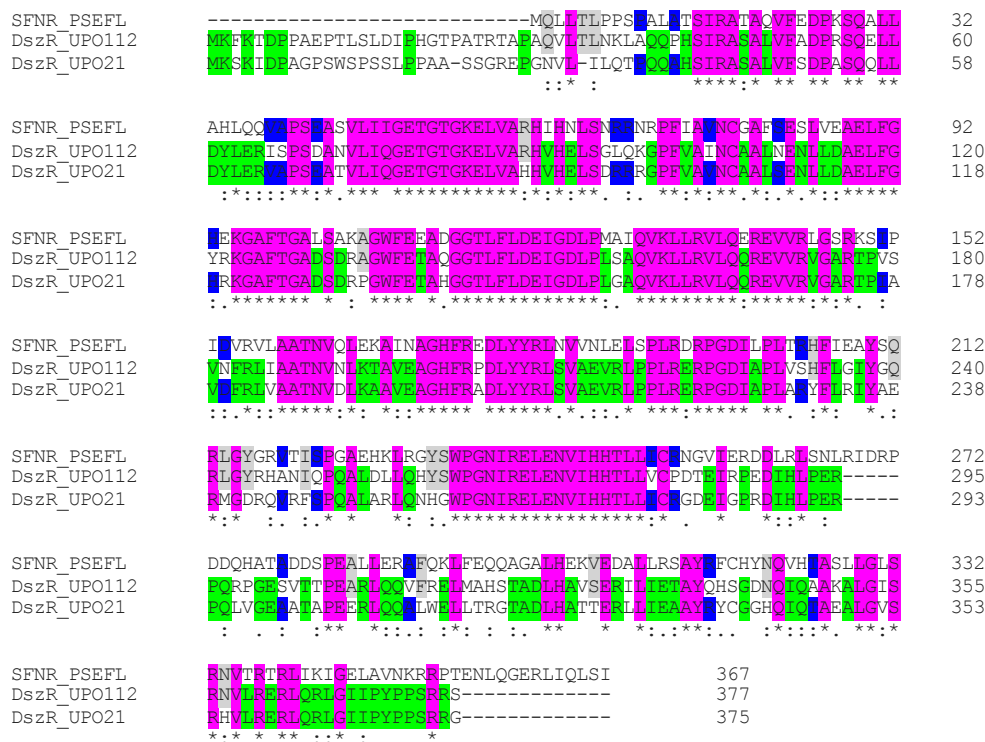


Figura 10: Alineamiento múltiple de SfnR de *P. Fluorescens* (Q3K999.1), DszR_{UPO112} y DszR_{UPO21}, realizado con clustalomega. Se resaltan en fucsia los residuos conservados entre las 3 secuencias, en verde los conservados entre DszR_{UPO112} y DszR_{UPO21}, en gris los conservados entre DszR_{UPO112} y SfnR y en azul los conservados entre DszR_{UPO21} y SfnR.

Se realizó un análisis filogenético de DszR de UPO21 y UPO112, junto con representantes de cada uno de los grupos de proteínas más relevantes, a los que mostraban similitud. En el árbol podemos observar (figura 11), que ambas están en la misma rama y muy cercanas a los SfnR, difiriendo bastante más con otros activadores de σ^{54} sin dominio N-terminal regulador.

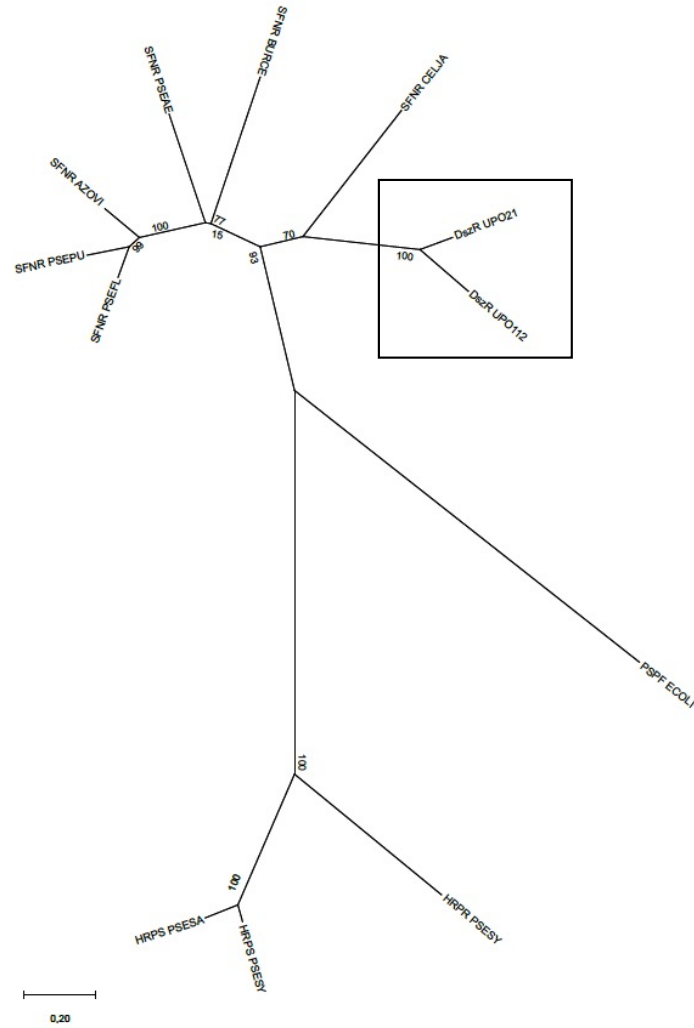


Figura 11: Árbol filogenético de DszR_{UPO112} y DszR_{UPO21} (se destacan con un recuadro estas 2 proteínas) junto con los siguientes activadores transcripcionales dependientes de σ^{54} , del grupo IV (Bush & Dixon, 2012): PSPF_ECOLI, activador transcripcional del operón Psp de *Escherichia coli* K-12 (P37344.2). HRPS_PSESY, locus de patogenicidad probable proteína reguladora HrpS de *Pseudomonas syringae* (P37931.1). HRPS_PSESA, locus de patogenicidad probable proteína reguladora HrpS de *Pseudomonas savastanoi* (P26316.2). HRPR_PSESY, locus de patogenicidad probable proteína reguladora HrpR de *Pseudomonas syringae* (P37930.1). SFNR_PSEFL, regulador transcripcional dependiente de σ^{54} SfnR de *Pseudomonas fluorescens* (Q3K999). SFNR_PSEPU, regulador transcripcional dependiente de σ^{54} SfnR de *Pseudomonas putida* (Q845S7). SFNR_PSEAE, posible regulador transcripcional dependiente de σ^{54} de *Pseudomonas aeruginosa* (A0A0H2ZA37). SFNR_AZOVI, proteína activadora, regulador transcripcional dependiente de σ^{54} de *Azotobacter vinelandii* (M9YI72). SFNR_BURCE, regulador transcripcional dependiente de σ^{54} de *Burkholderia cenocepacia* (B4EPG9). SFNR_CELJA, regulador transcripcional dependiente de σ^{54} de *Cellvibrio japonicus* (B3PJX0).

En cuanto a la proteína codificada por *dszH* de UPO112, tal y como se ha mencionado en el apartado I.2.1 de resultados, según la base de datos nr, la proteína a la que más se parece es una alfa/beta hidrolasa de *Propionivibrio dicarboxylicus* (WP_176785911.1) con la que tiene una identidad del 54,88%, con una cobertura del 75% e value 9e-50; sin embargo, no se parece a ninguna de las proteínas bien caracterizadas de la base de datos SwissProt. Además, *dszH* no parece tener homología con ningún gen del genoma de *Comamonadaceae bacterium* ni con el gen aguas arriba de *dszR* en UPO21 (*gen11*). (tablas 5 y 8).

A pesar de ello, y puesto que *dszH* y *dszR* están solapados en UPO112, aunque no en UPO21 (tabla 7), y puesto que dicho solapamiento sugiere acoplamiento traduccional, y de alguna forma una posible implicación en la degradación de DBT, se pensó que estos genes pudieran estar relacionados, por lo que se realizó un alineamiento de DszHUPO₁₁₂ y la proteína codificada por *gen11*_{UPO21} (alfa beta hidrolasa, QBM04638.1), comprobando que ambas presentan un 37,63 % de identidad, con una cobertura del 86% y 2e-32 de e value.

Al realizar un alineamiento múltiple con clustalomega entre la proteína codificada por *gen11*_{UPO21} (alfa beta hidrolasa, QBM04638.1), DszHUPO₁₁₂ y la alfa beta hidrolasa de *Propionivibrio dicarboxylicus*, según se muestra en la figura 12, comprobamos que DszH presenta un extremo amino más largo, y que exceptuando el extremo carboxilo, que es muy diferente entre las tres proteínas, los residuos conservados entre las tres secuencias se encuentran distribuidos de forma bastante homogénea a lo largo de ellas. Además, al analizar los dominios conservados en cada una de ellas con blastp, podemos concluir que las tres proteínas poseen un dominio conservado de la superfamilia alfa/beta hidrolasa, la cual incluye proteínas con una gran diversidad funcional, entre ellas hay proteasas, lipasas, peroxidasas, esterasas, epóxido hidrolasas y deshalogenasas.

Por otra parte; puesto que *dszH* solapa con *dszR*, y que DszR es muy parecido a SfnR de *Pseudomonas fluorescens*, y teniendo en cuenta que CysB se une de la región promotora de *sfnECR* en secuencias ricas en AT aunque sin secuencia consenso, (Kouzuma *et al.*, 2008) tal y como se ha comentado en la introducción, se ha analizado la región aguas arriba de *dszH*, encontrándose una región entre el -34 y el -112, con respecto al ATG de *dszH*, con un contenido en AT > 75%.

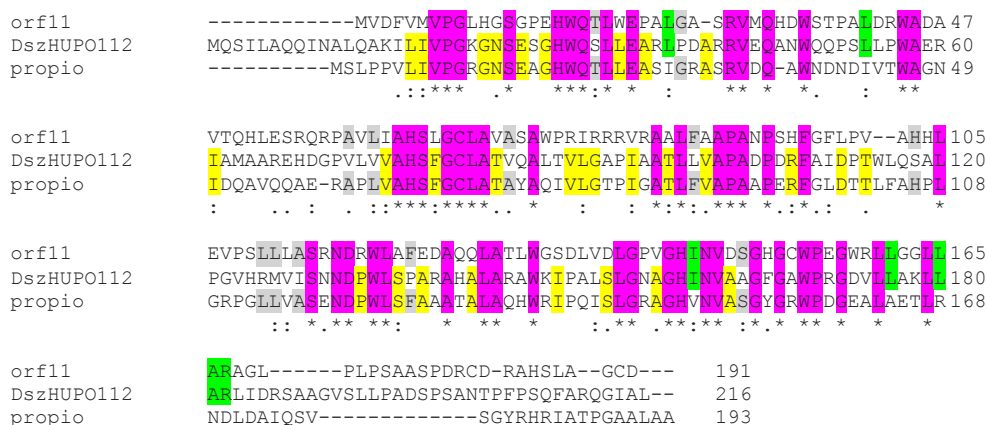


Figura 12: Alineamiento múltiple de la proteína codificada por *gen11*_{UPO21} (QBM04638.1) vs DszH_{UPO112} vs la alfa beta hidrolasa de *Propionovibrio dicarboxylicus* (WP_176785911.1) realizado con clustalomega. Se resaltan en color fucsia los residuos conservados en las 3 secuencias, en color verde los conservados entre DszH_{UPO112} y la proteína codificada por *gen11*_{UPO21}, en color amarillo los conservados entre DszH_{UPO112} y alfa beta hidrolasa de *Propionovibrio dicarboxylicus* y en color gris los conservados entre la proteína codificada por *gen11*_{UPO21} y la alfa beta hidrolasa *Propionovibrio dicarboxylicus*

1.2.3. Isla genómica.

Parece ser que tanto UPO21 como UPO112 contienen un fragmento de ADN delimitado por los genes *dsz*, tanto catabólicos con reguladores, que se parece mucho entre ambos (figura 7). Sin embargo, este fragmento se encuentra ubicado en contextos genómicos completamente diferentes puesto que las secuencias de los clones UPO21 y UPO112 que flanquean los genes *dsz* son muy diferentes entre sí, lo que sugiere que este fragmento común podría constituir una isla genómica. Esta isla, ligeramente mayor en UPO112, (10278 pb, y 9502 pb en UPO21), se habría insertado en eventos independientes y en distintas partes del genoma de distintas bacterias. Particularmente, en el caso de UPO112, de alguna estirpe de la especie *Candidatus* Macondimonas diazotrophica, ya que los genes flanqueantes de los genes *dsz* muestran la mayor similitud con genes de esta especie bacteriana.

Para apoyar la hipótesis de que los genes DBT se hayan insertado en el genoma de *Candidatus* Macondimonas diazotrophica, se comparó su estructura genómica, con la de las *orfs* de UPO112. Los resultados se muestran en la tabla 10 y podemos observar que desde la *orf1* a la 15 de UPO112, corresponden con *orfs* correlativas del genoma de *Candidatus* Macondimonas diazotrophica, (*orfs* 173 a 184), la *orf13* debe haber sufrido alguna reorganización entre un organismo y otro, y de nuevo las *orfs* 14 y 15 vuelven a ser correlativas, concordando con las *orfs* 185 y 186. Las *orf* 16 a 21, (salvo la *orf19*), se corresponden con las *orfs* correlativas 512-517. Después de estas *orf*, en

UPO112 se encuentran los genes con identidad con UPO21 que constituirían la isla genómica, y a continuación, hay una zona de transposasas, integrasas, ligasas, secuencias de inserción, y proteínas hipotéticas. Por último, hacia el final del inserto (*orfs39-42*), de nuevo hay homología con *Candidatus* Macondimonas diazotrophica, correlacionando con las *orfs* 1150 a 1153.

Considerando que los genes que flanquean la posible isla genómica en UPO112 también son correlativos en el genoma de *Candidatus* Macondimonas diazotrophica, y que en UPO112 se detectan genes completos o fragmentos génicos potencialmente codificantes de enzimas asociadas a ADN móvil (transposasas, integrasas) justo colindantes con el operón *dszEABC*, podemos suponer que es probable que la bacteria haya adquirido los genes de degradación de DBT de forma horizontal, habiéndose insertado en su genoma como una isla genómica.

| <i>orfs en UPO112</i> | Nº de <i>orf</i> del genoma de <i>Candidatus</i> Macondimonas diazotrophica |
|-----------------------|---|
| <i>orf1</i> * | 173 |
| <i>orf2</i> | 174 |
| <i>orf3</i> | 175 |
| <i>orf4</i> | 176 |
| <i>orf5</i> | 177 |
| <i>orf6</i> | 178 |
| <i>orf7</i> | 179 |
| <i>orf8</i> | 180 |
| <i>orf9</i> | 181 |
| <i>orf10</i> | 182 |
| <i>orf11</i> | 183 |
| <i>orf12</i> | 184 |
| <i>orf13</i> | 225 |
| <i>orf14</i> | 185 |
| <i>orf15</i> | 186 |
| <i>orf16</i> | 512 |
| <i>orf17</i> | 513 |
| <i>orf18</i> | 514 |
| <i>orf19</i> | 550 |
| <i>orf20</i> | 516 |
| <i>orf21</i> | 517 |
| <i>orf22 - dszH</i> | |
| <i>orf23 - dszR</i> | |
| <i>orf24</i> | |
| <i>orf25</i> | |
| <i>orf26</i> | |
| <i>orf27</i> | |
| <i>orf28 - dszC</i> | |
| <i>orf29 - dszB</i> | |
| <i>orf30 - dszA</i> | |
| <i>orf31 - dszE</i> | |
| <i>orf32*</i> | |
| <i>orf33</i> | |
| <i>orf34*</i> | |
| <i>orf35</i> | |
| <i>orf36</i> | |
| <i>orf37</i> | |
| <i>orf38*</i> | |
| <i>orf39</i> | 1150 |

RESULTADOS

| | |
|---------------|------|
| orf40 | 1151 |
| orf41 | 1152 |
| orf42* | 1153 |

Tabla 10: *orfs* de UPO112, se resaltan en negrita las *orf* con homología con *Candidatus* Macondimonas diazotrophica. (* indica proteína incompleta)

II. Caracterización de la regulación del operón.

II.1. Caracterización de la expresión del operón *dszEABC* de UPO21.

II.1.1. Expresión propia y desde P_{sal} . Regulación postranscripcional.

En el trabajo de Tesis de la Dra. Laura Terrón, se construyeron una serie de fusiones traduccionales a *lacZ* y transcripcionales a *gfp* de los genes DBT de UPO21, para iniciar el estudio de la expresión del operón *dszEABC* y se concluyó que estos genes deben tener algún mecanismo de regulación postranscripcional, que impide su traducción desde transcritos espúreos generados por promotores diferentes al propio.

Para complementar los resultados del citado trabajo y caracterizar con más detalle la expresión de dicho operón; en el presente trabajo se han construido fusiones transcripcionales al gen reportero *lacZ* de los genes *dszE*_{UPO21} (pMPO1395), *dszEA*_{UPO21} (pMPO1390) y *dszEABC*_{UPO21} (pMPO1388), y se ha construido una nueva fusión traduccional *dszE*_{UPO21}::*lacZ* (pMPO1394), realizada en el nucleótido 262, a diferencia de la que se empleó en la Tesis de la Dra. Terrón que era una fusión al codón de inicio, ATG. Todas estas fusiones se construyeron en el vector pMPO579, para poder aumentar el número de copias de los plásmidos con arabinosa, y tratar de inducir su expresión desde el promotor heterólogo P_{sal} . Los detalles de las construcciones se encuentran en la sección materiales y métodos, pero se representan esquemáticamente en la figura 13.

Con estas fusiones se procedió a caracterizar en profundidad la expresión de los diferentes genes *dsz* tanto desde su propio promotor, para lo cual se transfirieron a la estirpe EPI300-T1^R, como desde el promotor heterólogo, transfiriéndose en este caso a la estirpe MPO554 NaI^R, e induciendo con salicilato. El gen *dszR* se aportó mediante el plásmido pMPO588, y se sustituyó por el vector vacío (pIZ1016) en los casos en que no se deseaba aportar DszR. Las distintas estirpes con los distintos plásmidos se indujeron en medio mínimo con DBT, arabinosa, salicilato e IPTG, tal y como se describe en materiales y métodos, y se ensayó su actividad β -galactosidasa tras 6 horas de inducción. Tal y como se observa en la figura, los valores alcanzados por las fusiones transcripcionales son aproximadamente 10 veces superiores a los alcanzados por las traduccionales, y son similares si comparamos las tres fusiones entre sí, indicando una coregulación de los tres genes.

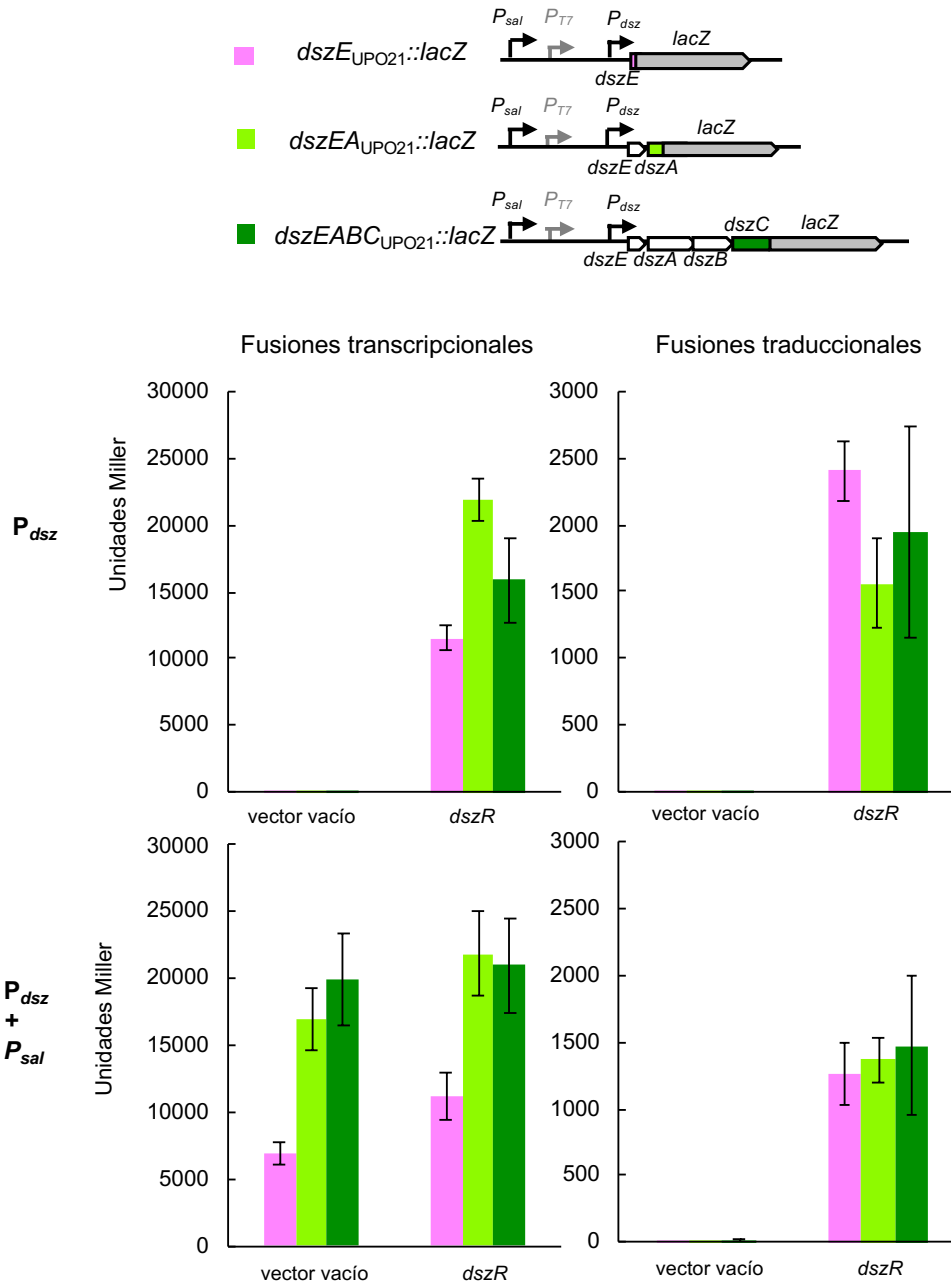


Figura 13: Panel superior: esquema de las construcciones de las fusiones transcripcionales al gen indicador *lacZ* de los genes *dszE_{UPO21}* (pMPO1395), *dszEA_{UPO21}* (pMPO1390) y *dszEABC_{UPO21}* (pMPO1388), y traduccionales *dszE_{UPO21}::lacZ* (pMPO1394), *dszEA_{UPO21}::lacZ* (pMPO597) y *dszEABC_{UPO21}::lacZ* (pMPO593). **Panel inferior:** Ensayo de actividad betagalactosidasa de las fusiones transcripcionales (paneles de la izquierda), y transcripcionales (paneles de la derecha) a los genes *dszE_{UPO21}*, *dszA_{UPO21}*, y *dszC_{UPO21}*. Expresión desde su propio promotor (paneles superiores), y expresión desde su propio promotor sumada a la expresión desde P_{sal} (paneles inferiores). En cada gráfica se muestran los valores con la incorporación del plásmido pMPO588 (que porta *dszR*) (columnas de la derecha en cada gráfica) o con el correspondiente

plásmido vacío, plZ1016, (columnas de la izquierda en cada gráfica). Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

Por otra parte, la expresión de *dszE*, *dszEA* y *dszEABC*, desde su propio promotor, tanto a nivel transcripcional, como traduccional, es estrictamente dependiente de la presencia de DszR, lo que demuestra su implicación en la expresión del operón. Cuando las fusiones se expresan en MPO554 NaI^R , y se inducen con salicilato, observamos por tanto el efecto sumado de la expresión desde P_{sal} y la expresión desde P_{dsz} . En este caso, comprobamos que las 3 fusiones pueden transcribirse, aunque no se aporte DszR, como cabe esperar de una transcripción heteróloga desde P_{sal} . Sin embargo, sorprendentemente, las fusiones traduccionales a los genes *dsz* solo se expresaron en presencia de DszR. Esto hace suponer que la transcripción desde P_{sal} , aguas arriba del promotor propio, no es productiva, lo que sugiere que debe existir un mecanismo de regulación postranscripcional que sería responsable del distinto comportamiento de las fusiones transcripcionales y traduccionales.

Las diferencias observadas en la traducibilidad pueden ser debidas a la presencia de DszR, lo que indicaría que además de ser un activador transcripcional, jugaría un papel importante en la traducción de los genes del operón. Pero, por otra parte, el hecho de que los transcritos provenientes del promotor heterólogo no puedan traducirse, podría deberse a que la secuencia leader es diferente según dónde se inicie la transcripción. Puesto que la secuencia leader del ARNm podría tener importancia en la regulación de la traducción, por ejemplo, conteniendo un posible sitio de unión de un represor que ocluyera el RBS, (Samal, 2013) o una estructura secundaria que impidiese la traducción del ARNm o redujese su estabilidad, una región 5' distinta podría claramente afectar la funcionalidad del ARNm.

II.1.2. Determinación del inicio de transcripción desde P_{dsz} .

Para profundizar en la posible implicación de la región 5' en la traducibilidad de *dszE*, se estudió en detalle la región aguas arriba de la secuencia codificante de dicho gen, tanto de UPO21 como de UPO112, y se encontró un posible sitio de promotores dependientes de σ^N (figura 14).

| | | |
|--------|--|-----|
| UPO21 | TTCGGCTGAAAAGCGCATTAACCATTCAAATACAGTGTGTTGAATCCTGGCATGACCTCT | 118 |
| UPO112 | TAAAAACGACCGATAAAATCCCAAATAAAACAA--CTAGATATAAATGGCATAAGAAAT | 118 |
| | * ** *** * * * * * * * * * ***** * | |
| UPO21 | GCTTTATGCATATTGGATCTGAAAAATCCAGTTGTTTCAT--AAGGAAACGTCATGGCAAA | 176 |
| UPO112 | GCTTTTATAGGCTGTGTCTTTGCATATACAGCTCTTAAATAAAGGAAGTGCTATGCCTCA | 178 |
| | ***** * ** * * * * * * * * * ***** * *** * * | |

Figura 14: Alineamiento de la región promotora de *dszE*_{UPO21} y *dszE*_{UPO112}. Se resalta en gris el posible promotor dependiente de sigma54, y en fucsia el codón de inicio de *dszE*.

Para confirmar su implicación en la transcripción de los genes *dsz*; se realizó un experimento de determinación del 5', con el oligonucleótido PEX-DBT1, tal y como se detalla en materiales y métodos. El experimento se realizó tanto con la fusión transcripcional, como con la traduccional *dszEA_{UP021}::lacZ*, obteniendo resultados similares. Tal y como se muestra en la figura 15, se detectó un inicio de la transcripción, en una G situada 13 pb aguas abajo del posible promotor (GG-N₁₀-GC), y 35 pb aguas arriba del ATG de *dszE*. El experimento se realizó con ARNm procedente de cultivos con *dszR* y sin *dszR* (datos no mostrados) pero solo se detectó señal en presencia de *dszR*.

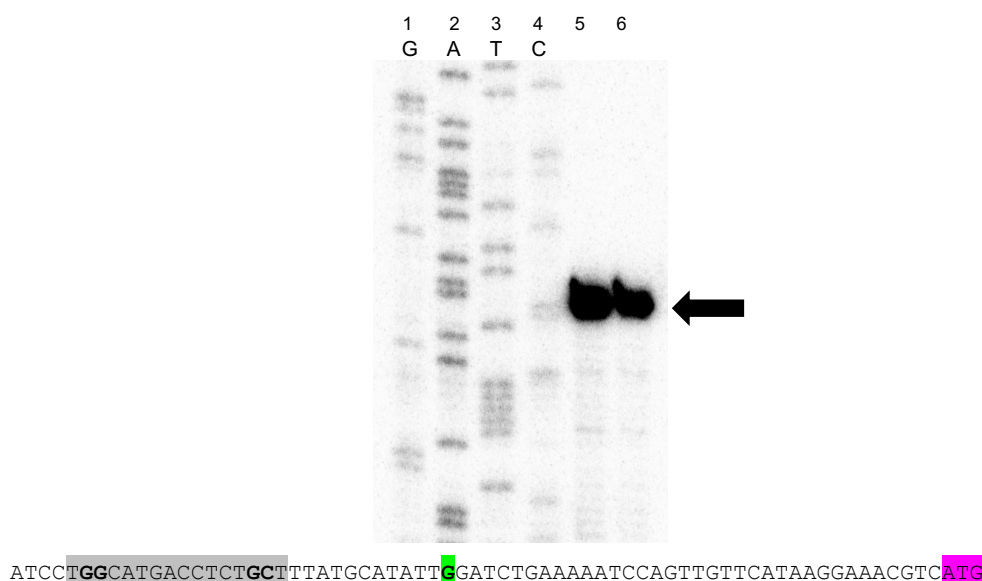


Figura 15: Determinación del inicio de transcripción de *dszE*. **Panel superior:** Reacciones de extensión desde el cebador. Carriles 1 al 4: GATC, reacciones de secuenciación para localizar el inicio de transcripción; carril 5: ADN retrotranscrito de la fusión transcripcional *dszEA_{UP021}::lacZ*, carril 6: ADN retrotranscrito de la fusión traduccional *dszEA_{UP021}::lacZ*. La flecha indica el inicio de transcripción. **Panel inferior:** Secuencia de la región promotora de *dszE_{UP021}*. Se resalta en gris el posible promotor dependiente de sigma54, en fucsia el codón de inicio de *dszE* y en verde (y negrita) el inicio de transcripción determinado experimentalmente.

Se realizaron otros experimentos del mismo tipo tratando de detectar un inicio de transcripción de *dszA* que pudiera ser independiente al de *dszE*, pero no se obtuvo ningún resultado lo cual apoya fuertemente la idea de que ambos formen parte del mismo operón, y que no haya un operón interno en *dszE* desde el que se puedan transcribir los genes aguas abajo (no mostrado).

Por otra parte, también se trataron de detectar posibles ARNs antisentido, que pudieran tener algún papel regulador, pero no se obtuvo ninguna señal (no mostrado).

II.1.3. Mutantes de delección ΔP_{dsz} . Expresión desde P_{sal} .

Tras la determinación del inicio de la transcripción, se diseñaron fusiones traduccionales a $dszE_{UPO21}$ (pMPO1396) y $dszEA_{UPO21}$ (pMPO1541), cuya construcción se detalla en materiales y métodos y se esquematiza en la figura 16, en las que se deleccionaba la secuencia aguas arriba del inicio de la transcripción y se sustituía el P_{dsz} por el promotor heterólogo P_{sal} , que quedaba así posicionado en el mismo sitio en el que originalmente se encuentra el promotor propio, con el fin de producir un transcrito con la misma secuencia que el transcrito nativo, pero independiente de DszR.

Las fusiones se testaron en EPI300-T1^R y MPO554 NaI^R, aportando $dszR$ en el plásmido pMPO588 y utilizando como control el vector vacío, pIZ1016, en medio mínimo con DBT, arabinosa, salicilato e IPTG, tal y como se indica en materiales y métodos. En la figura 16 se muestran los resultados, en comparación con los obtenidos en el caso de las fusiones silvestres, para una mejor comprensión de los mismos.

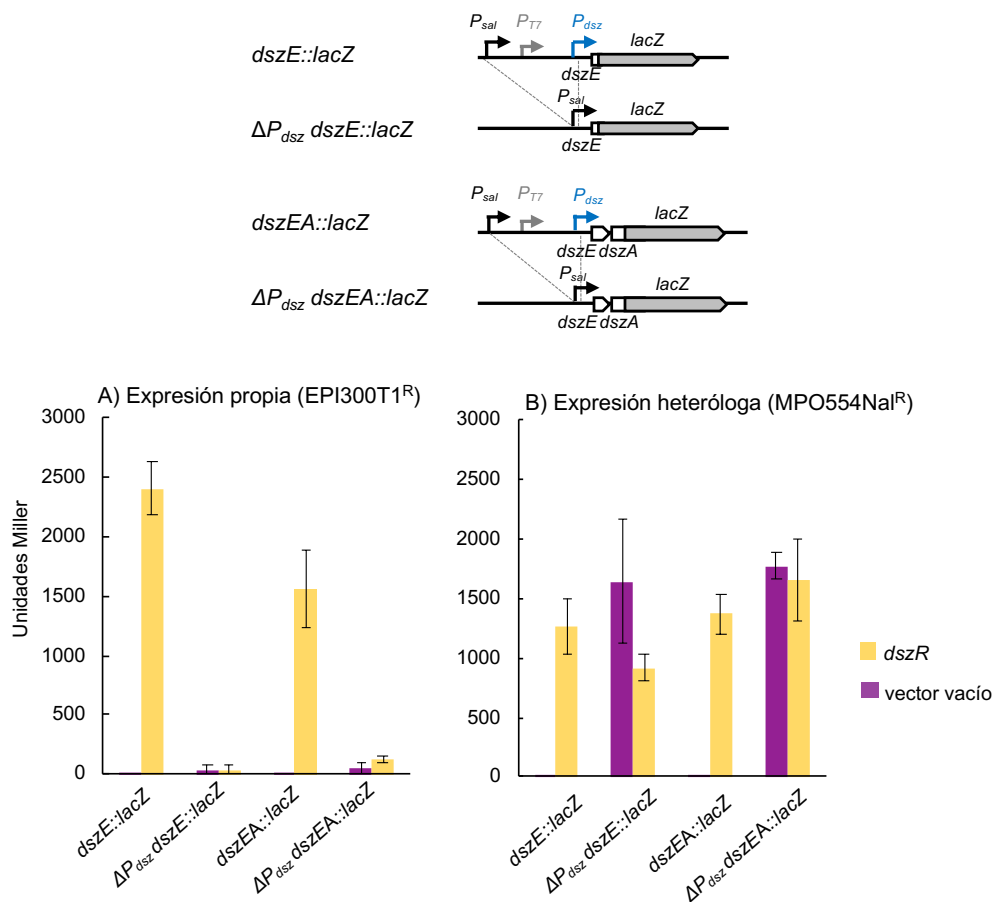


Figura 16: Delección de P_{dsz} . **Panel superior:** esquema de las construcciones de las fusiones traduccionales al gen indicador $lacZ$ que incluyen la delección del promotor de los genes dsz , de forma que P_{sal} , resulta reposicionado en el lugar que originalmente ocupa P_{dsz} . Fusión a $dszE_{UPO21}$ (pMPO1396) y a $dszEA_{UPO21}$ (pMPO1541). **Panel inferior:** Ensayo de actividad

betagalactosidasa de las fusiones traduccionales silvestres y con el promotor *dsz* deletado, en EPI300, donde se observa la expresión propia (panel A), y la expresión heteróloga desde P_{sal} en MPO554NaI^R al inducir con salicilato (panel B). En cada gráfica se muestran los valores con la incorporación del plásmido pMPO588 (que porta *dszR*) y con el correspondiente plásmido vacío (pIZ1016). Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

Cuando las fusiones se ensayaron en la estirpe EPI300-T1^R (figura 16, panel A), se observa que las fusiones mutantes, que contienen la delección solo pueden expresarse a un nivel basal muy bajo, ya que carecen de su propio promotor, y P_{sal} no puede inducirse en esta estirpe. Sin embargo, en la estirpe MPO554 NaI^R en la que los genes se pueden transcribir desde P_{sal} , se observa que la expresión en ambos mutantes de delección es elevada; siendo, además, independiente de DszR. Este resultado muestra que, si el operón *dszEABC* se transcribe desde su inicio de transcripción original, aunque sea un promotor distinto al suyo, el ARNm resultante puede traducirse perfectamente.

II.1.4. Estructura secundaria del ARNm.

Se realizó un análisis de la estructura de los dos posibles ARNm de UPO21 con la herramienta RNAfold Web Server (Lorenz *et al.*, 2011), desde el sitio de inicio de la transcripción nativo y desde 35 nucleótidos aguas arriba (e incluyendo en ambos casos los 30 primeros codones de *dszE*). Como podemos observar en la figura 17, cuando la transcripción se inicia desde el propio promotor, la estructura de ARNm, muestra distintas horquillas con bucles en las que se observa que la SD no forma parte de una estructura secundaria y el codón de inicio estaría solo parcialmente oculto en una estructura secundaria bastante débil por lo que el sitio de unión de ribosomas estaría razonablemente accesible. Sin embargo, si la transcripción se produce desde aguas arriba, se observan dos grandes estructuras secundarias en la más estable de las cuales, tanto el codón de inicio de *dszE*, como la SD se encuentran ocultos, por lo que el sitio de unión de ribosomas estaría mucho menos accesible. En la formación de estas estructuras secundarias está claramente implicada la secuencia promotora dependiente de sigma N. Comparando la traducibilidad de ambos transcritos con la herramienta RBS calculator (Salis *et al.*, 2009), observamos que la traducibilidad es ≈ 47 veces mayor en el caso de los transcritos desde P_{dsz} (de 1446 ua a 31 ua). Análisis similares pero incluyendo secuencias que se extendían paulatinamente hasta 450 pb aguas arriba de los genes *dsz*, mostraron que tanto las estructuras secundarias ocluyentes del sitio de unión de ribosomas como la baja traducibilidad de estos ARN mensajeros que se

transcriben desde aguas arriba del inicio de transcripción nativo, se seguían manteniendo, lo que indica claramente que *dszE* no puede traducirse si el transcrito proviene de aguas arriba y contiene la secuencia promotora nativa transcrita en la región líder de ese ARNm.

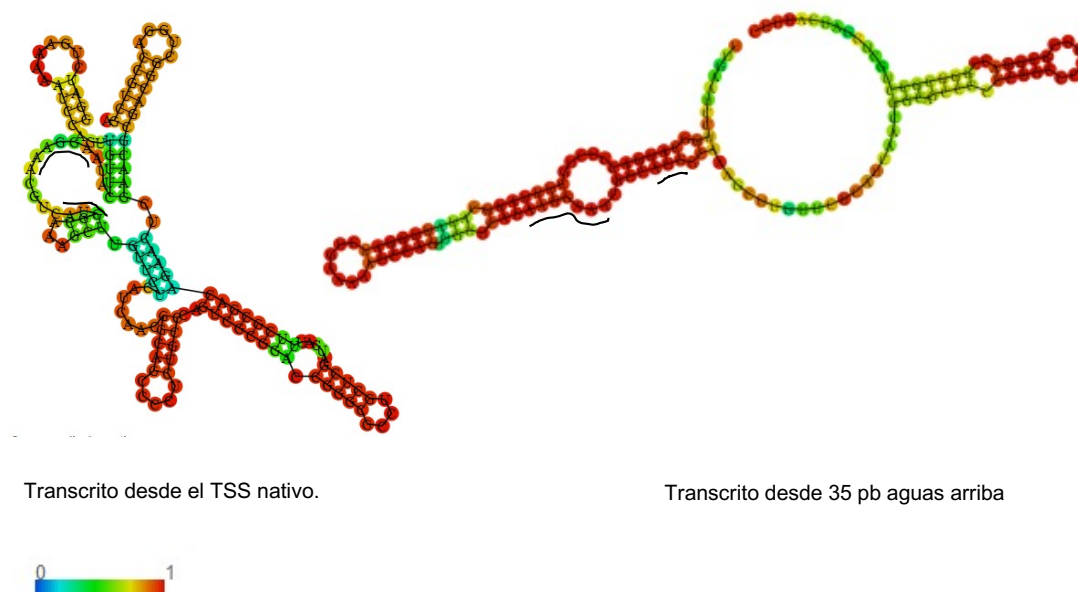


Figura 17: Estructuras secundarias predichas de la región líder de *dszE*. **Izquierda:** Transcrito desde el TSS nativo. **Derecha:** Transcrito desde 35 pb aguas arriba. Se marcan con una línea el codón de inicio de *dszE* y la secuencia Shine-Dalgarno. Las estructuras propuestas son dinámicas y la escala de color indica la frecuencia de cada nucleótido de estar en la situación propuesta en la estructura, siendo azul la opción más frecuente, y el rojo la más frecuente.

II.2. Caracterización del papel de DszR.

Una vez definida la expresión del operon *dszEABC* solo si se transcribe desde su propio promotor, tratamos de caracterizar con más detalle la implicación de DszR en la expresión del operón.

II.2.1. Actividad de DszR *in vivo*.

Como se ha mencionado en el apartado I.2 de resultados; DszR codificado en el fósido UPO21, fue inicialmente descrito como un activador transcripcional de promotores dependientes de σ^N (Terrón, 2014), aunque en la Tesis de la Dra. Terrón no quedaba claro su mecanismo de funcionamiento. Con anterioridad a dicho trabajo, se habían asociado reguladores dependientes de σ^{54} con la utilización de DBT como fuente de carbono y energía (Di Gregorio *et al.*, 2004), pero no con ninguna ruta de desulfuración de DBT. Tal y como se menciona en el apartado I.2.2 de resultados, DszR carece del dominio regulador N-terminal por lo que podría formar parte del grupo IV de

activadores transcripcionales, que son constitutivamente activos, siendo su actividad controlable por interacción con otra proteína sensora (Bush & Dixon, 2012).

Para confirmar o descartar la hipótesis de que la actividad de DszR *in vivo* fuera constitutiva se realizaron ensayos de expresión de la fusión transcripcional *dszE_{UPO21}::lacZ* (pMPO1395), en la estirpe EPI300-T1^R, aportando DszR_{UPO21} en el plásmido pMPO588 (o sustituyéndolo por el vector vacío pLZ1016). La fusión se indujo en medido mínimo con diferentes fuentes de azufre: sulfato (para comprobar si hay inducción cuando hay una fuente rica de azufre), DBT (para comprobar que la fusión se induce de forma correcta), sulfato y DBT simultáneamente (para testar si existía represión por fuente preferencial de azufre) o sin azufre (pero aportando NNDMF, que es el solvente del DBT, por si causara algún efecto). La expresión se testó tras 24 horas de inducción, sin y con la presencia de IPTG.

Los resultados que se muestran en la figura 18, confirman que la expresión de *dszE* es dependiente de DszR, ya que, en su ausencia, se obtienen niveles basales de expresión. En su presencia, aún sin IPTG para inducir la transcripción de su gen codificante, se observan unos niveles de expresión considerables, que aumentan incluso más en presencia de IPTG. Tanto en ausencia como en presencia de IPTG, los niveles de expresión de *dszE::lacZ* no dependen de la fuente de azufre, ya que los niveles son prácticamente iguales en todas las fuentes de azufre testadas y las pequeñas diferencias que se observan entre ellas, sobre todo en presencia de IPTG, son achacables a las diferencias en su estado metabólico, observándose una mayor expresión en los cultivos que pueden crecer al contener sulfato como fuente de azufre.

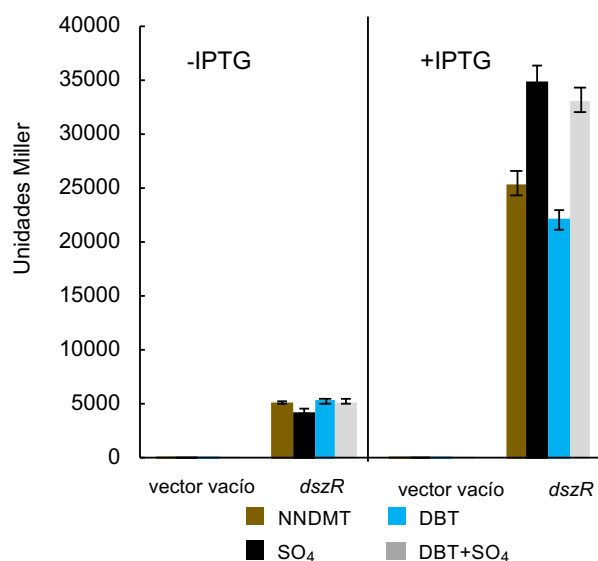


Figura 18: Actividad de DszR *in vivo*. Ensayo de actividad betagalactosidasa de la fusión transcripcional *dszE_{UPO21}::lacZ* (pMPO1395), en la estirpe EPI300-T1^R, aportando DszR_{UPO21} en

el plásmido pMPO588 (columnas de la derecha) o sustituyéndolo por el vector vacío pIZ1016 (columnas de la izquierda), en medio mínimo con diferentes fuentes de azufre, sin inducir (gráficas de la izquierda) o induciendo con IPTG (gráficas de la derecha).

II.2.2. Caracterización *in vitro* de la actividad de DszR.

Una vez descrita la actividad de DszR *in vivo*, se decidió caracterizar su actividad *in vitro*.

II.2.2.a. Purificación de DszR.

Para poder abordar la caracterización de la actividad *in vitro* de DszR, era necesario purificar la proteína, lo cual se realizó empleando el sistema IMPACT™ de New England Biolabs que permite obtener una proteína totalmente nativa, ya que la inteína que se emplea, se escinde de la proteína sin dejar cicatriz.

El protocolo descrito por la casa comercial se adaptó, y tras varios intentos finalmente se consiguió purificar DszR con una pureza estimada del 98% (tal y como se observa en la figura 19) y una concentración de 5,1 μ M, asumiendo que DszR es un dímero.

Para ello se empleó el plásmido pMPO1544 (*dszR*_{UP021} clonado en el vector pTYB21, que fusiona la inteína en el extremo amino terminal de la proteína), cuya construcción se detalla en materiales y métodos. Dicho plásmido se transformó en la estirpe superproductora NCM631, previamente transformada con pIZ227, y se indujo la sobreproducción de la proteína con IPTG, y en frío, tal y como se indica en materiales y métodos.

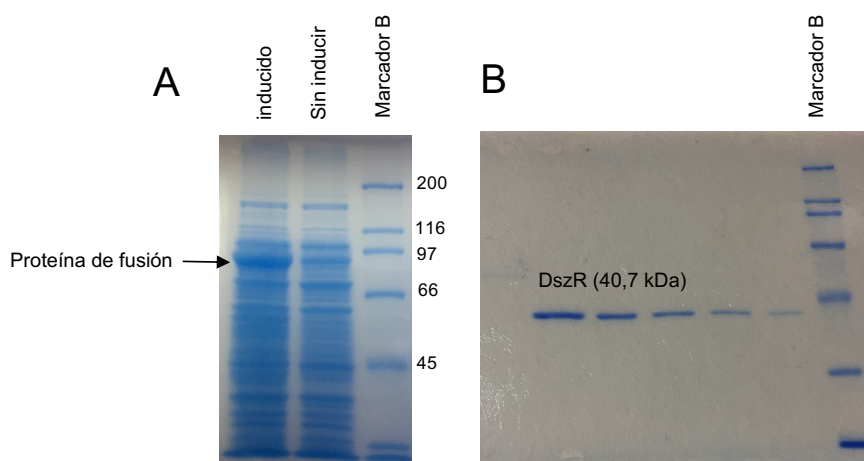


Figura 19: Purificación de DszR. **A)** gel de la superproducción de la proteína de fusión (DszR-inteína). Carril 1: cultivo inducido (se señala la banda de la proteína de fusión), carril 2: cultivo sin inducir, carril 3: marcador B (BioRad). **B)** gel de diferentes fracciones de la proteína DszR purificada (carriles 1 al 5), carril 6: marcador B (BioRad).

En la figura, se puede observar que la producción de la proteína de fusión se indujo correctamente; ya que se muestra la aparición de una banda 96,7KDa, resultado de la fusión de DszR (40,7 KDa) y la inteína de pTYB21 (56 KDa).

Una vez obtenido el pellet de la superproducción, se procedió a romper las células por sonicación, se centrifugó el extracto y se cargó el sobrenadante en una columna de quitina previamente equilibrada, como se indica en M&M. En este punto del proceso, hubo muchos problemas, puesto que la proteína de fusión no se quedaba retenida en la resina. Así que se decidió hacer varios pases por la columna con la mitad de la resina con un flujo $< 0,1$ ml/min e incubar el *flow through* del paso anterior, toda la noche con la otra mitad de la resina en agitación. Después, se empaquetó todo, y se indujo el corte de la inteína con DTT. Con este procedimiento se consiguió obtener DszR en varias fracciones de gran pureza, que se unieron para ser sometidas a diálisis con el fin de reducir la concentración de DTT. La concentración máxima que se pudo conseguir fue de $5,1 \mu\text{M}$; por encima de ésta, la proteína precipita.

II.2.2.b. Ensayos de retardo de movilidad en gel.

Para comprobar que la preparación de DszR obtenida tenía la capacidad de unirse a la región promotora de *dszE*, se realizaron ensayos de retardo en gel (EMSA), con una sonda obtenida por restricción, a partir de pMPO1547 y cantidades crecientes de DszR, tal y como se describe en M&M.

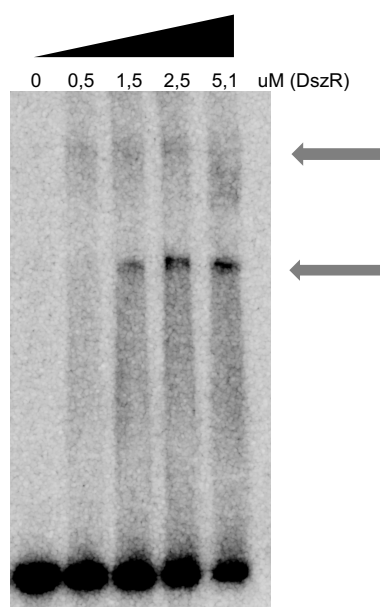


Figura 20: Retardo en gel de una sonda que contiene los posibles sitios de unión de DszR y P_{dsz} de UPO21 con cantidades crecientes de DszR. Se señalan 2 bandas retardadas cuya intensidad aumenta al aumentar la concentración de DszR.

Los resultados se muestran en la figura 20, y permiten apreciar, la desaparición de la sonda libre y la aparición de 2 bandas con movilidad relativa menor, cuya proporción con respecto a la sonda libre, aumenta al incrementar la concentración de DszR en la reacción. Con lo cual se confirmó que la preparación de DszR era activa y se decidió diseñar experimentos de protección frente a DNasaI.

II.2.2.c. Identificación de los sitios de unión de DszR.

Tras haber comprobado experimentalmente que DszR tenía la capacidad de unirse a la región reguladora del promotor *dszE_{UPO21}*, se analizó dicha región, con el objetivo de detectar posibles sitios de unión para la proteína activadora; encontrándose 2 sitios palindrómicos TGCN₁₀GCA separados 31 pb entre sí, y situados entre las posiciones -127 a -112 y -96 a -81, respecto al inicio de transcripción detectado experimentalmente. La comparación de las regiones promotoras-reguladoras de *dszE_{UPO21}* y *dszE_{UPO112}* (figura 21.A) muestra que el posible promotor (GG-N₁₀-GC), dependiente de σ^{54} se haya conservado en ambos, conservándose 11 de los 16 nucleótidos, así como su posicionamiento (-12/-24) con respecto al inicio de la transcripción, que coincide con lo habitual en los promotores reconocidos por los factores σ^{54} , que se ha definido experimentalmente en UPO21 (apartado II.1.2 de resultados) y se asume igual en UPO112. Los sitios palindrómicos detectados en *dszE_{UPO21}* también aparecen bien conservados en *dszE_{UPO112}* puesto que, tal y como se muestra en la figura 21.A, no sólo se conserva el patrón TGCN₁₀GCA y su separación de 31 pb, sino que también, presentan un alto grado de similitud en los nucleótidos adyacentes, ya que se conservan 7 de 18 en el sitio distal y 13 de 18 en el sitio proximal. Por último, se observa que la secuencia SD, también está totalmente conservada entre ambas regiones y que la principal diferencia entre ambas regiones se encuentra en la longitud de la región líder no traducida, que en UPO112 es de 37 pb, mientras que en UPO21, es de 35 pb. Estos datos permiten suponer que la expresión de los operones *dszEABC* de ambos fósidos es similar y que, probablemente, los sitios palindrómicos conservados sean los sitios de unión del activador DszR.

Por tanto, para corroborar que los posibles sitios de unión anteriormente descritos eran sitios de unión reales, se realizaron experimentos de protección frente a DNasaI. Para ello, se empleó una sonda que contiene el promotor de *dszE_{UPO21}* y secuencia aguas arriba, entre las posiciones -150 y +7, que se obtuvo por restricción a partir de pMPO1547, digiriendo con *EcoRV* y *Sall*, para marcar la cadena+ y con *HindIII* y *ApaI* para marcar la cadena -. (Los detalles, tanto de la construcción de dicho plásmido como de la realización del experimento se encuentran en M&M).

En la figura 21.B, se muestra el footprint de la cadena+, pudiendo observarse 2 ventanas de protección, cada una de las cuales comprende uno de los sitios conservados de unión predichos. Por tanto, se confirma que DszR se une a la región reguladora del promotor de *dszE* y que las regiones conservadas aguas arriba constituyen los sitios de unión de DszR.

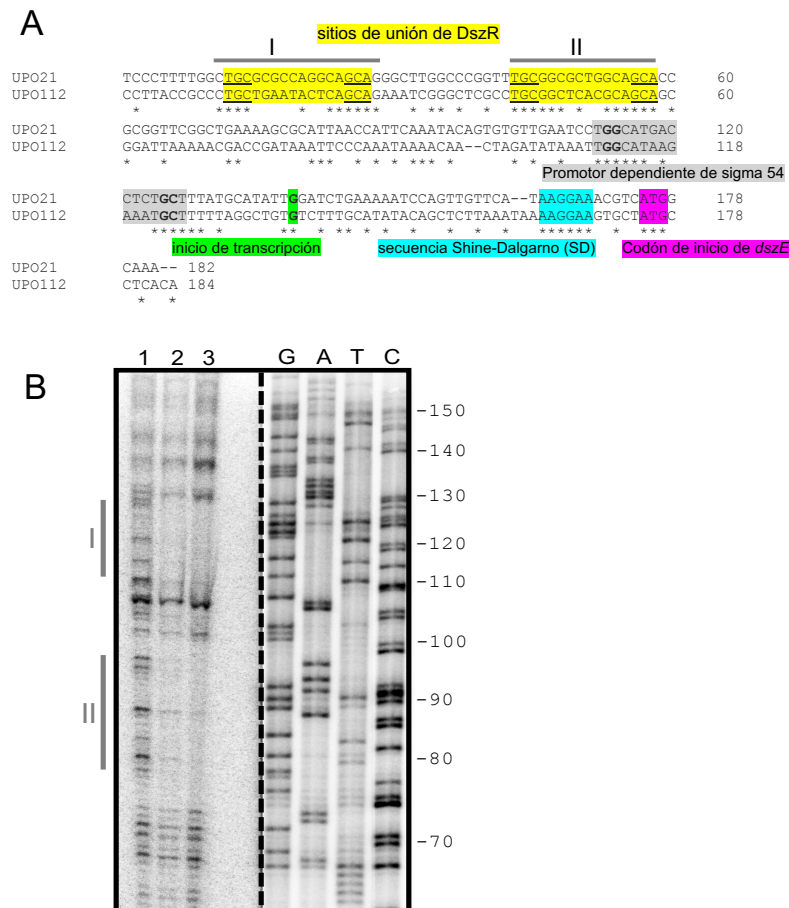


Figura 21: A) Alineamiento de la región promotora de *dszE*_{UPO21} y *dszE*_{UPO112}. Se marcan en amarillo los posibles sitios de unión de DszR, en gris el promotor dependiente de sigma54, en verde y negrita el inicio de la transcripción, en azul la secuencia Shine-Dalgarno (SD) y en fucsia el codón de inicio de *dszE*. Sobre la secuencia, se marcan con 2 líneas las zonas protegidas en el ensayo de protección frente a DNasal (I y II). **B)** Ensayo de protección frente a DNasal con una sonda que contiene la región de -150 a +7, con respecto al inicio de transcripción de *dszE*, y cantidades crecientes de DszR (se muestra la cadena+). Se señalan dos ventanas de protección (I y II). Carril 1: sin DszR, carril 2: 2,5 μ M DszR, carril 3: 5,1 μ M DszR, carriles 4 al 6: GATC, reacciones de secuenciación. Las imágenes de la secuenciación y de las reacciones de footprint provienen del mismo gel, pero con diferente grado de contraste, para una mejor visualización de los resultados.

II.2.2.d. Ensayos de Transcripción *In Vitro*.

Una vez demostrada la unión de DszR a sus secuencias predichas; se decidió realizar experimentos de Transcripción *in vitro*, para comprobar si efectivamente DszR activa la transcripción desde el promotor de *dszE*.

Se empleó como molde el plásmido pMPO1548, que contiene el $P_{dszUPQ21}$ (posiciones de -160 a +42), con el que se generaría un transcrito de 342 pb, 300 de las cuales son del plásmido. Con este ensayo también se pretendió comprobar si la transcripción de *dszE* era dependiente de IHF, tal y como está descrito para muchos otros promotores dependientes de σ^{54} (Shingler, 2011), ya que, además, como se observa en la figura 21.A, en la región reguladora de *dszE*, hay zonas ricas en AT aguas arriba de la secuencia promotora, las cuales están descritas como secuencias de unión de IHF, aunque la mala conservación de la secuencia consenso en los sitios de unión de esta proteína impide una identificación inequívoca de los mismos (Platero *et al.* 2012).

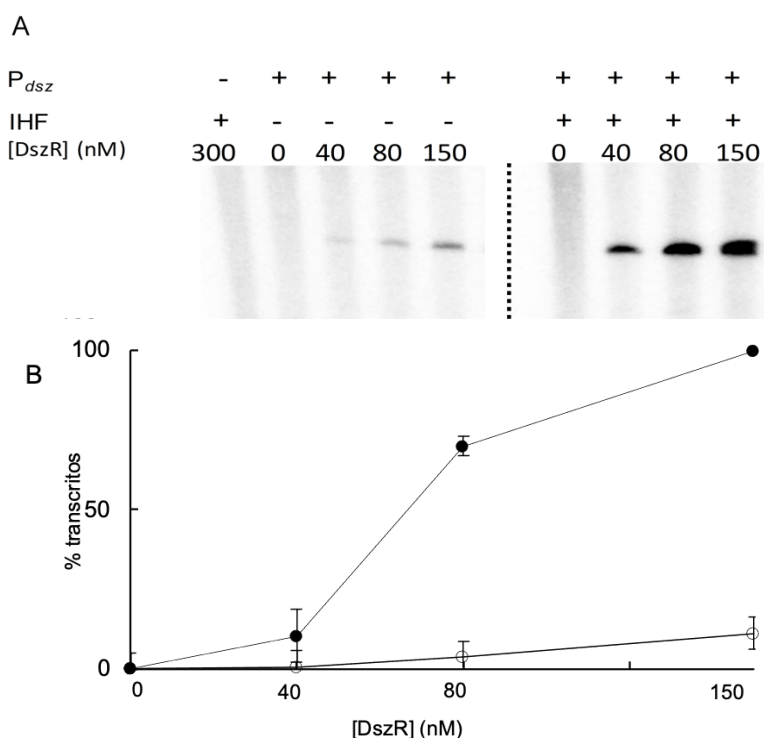


Figura 22: Transcripción *in vitro* desde P_{dsz} . **A:** Resultado primario de un experimento representativo de la transcripción *in vitro* desde el promotor *dsz*, con concentraciones crecientes de DszR, en presencia de σ^{54} y con la adición (derecha de la imagen) o no (izquierda de la imagen) de IHF. La línea discontinua indica que son imágenes que provienen de diferentes partes del mismo gel. **B:** Cuantificación de la transcripción, considerando 100%, el valor máximo alcanzado con la mayor cantidad de DszR (150 nM), los niveles de transcripción son relativos a dicho valor. Círculos vacíos: reacciones sin adición de IHF, círculos rellenos: reacciones con

adición de IHF. La cuantificación corresponde a la media de 3 experimentos y las barras de error representan la desviación estándar.

Tal y como se aprecia en la figura 22, en ausencia de DszR o de P_{dsz} , no es posible detectar transcripción. Sin embargo, en presencia de ambos se observa una banda de transcripción, cuya intensidad incrementa a medida que se aumenta la concentración de DszR en la reacción, con lo que queda demostrada la activación de la transcripción desde P_{dsz} por parte de DszR. Además, aunque en ausencia de IHF se detecta cierta transcripción, en presencia de dicha proteína la transcripción aumenta de forma sustancial: Si consideramos el máximo de transcripción obtenida en presencia de IHF como el 100%, en las mismas condiciones, pero sin IHF, la máxima transcripción obtenida es del 6%, confirmando así la dependencia de IHF para obtener una eficiente transcripción.

II.3. Regulación por fuente de azufre.

II.3.1. Expresión de $dszE_{UPO112}::lacZ$.

Teniendo en cuenta que las regiones reguladoras de $dszE$ son muy parecidas, así como la similitud que muestran las proteínas activadoras DszR, en UPO21 y en UPO112 (apartados I.2.2 y II.2.2.c de resultados), cabe suponer que la regulación del operón $dszEABC$, puede ser similar en ambos fósmidos. Sin embargo, es importante tener en cuenta que UPO112, a diferencia de UPO21, permite a *E. coli* crecer en DBT como única fuente de azufre, sin inducir el promotor P_{sal} , ni aumentar el número de copias del plásmido. Es decir, en UPO112 el operón $dszEABC$, puede expresarse por sí mismo, y es posible que esta expresión esté regulada por azufre. Para explorar esta posibilidad se decidió estudiar más en detalle la expresión de $dszE_{UPO112}$ y con este propósito, se construyó una fusión transcripcional al gen reportero $lacZ$, obteniéndose el plásmido pMPO1550, con el cual se procedió al estudio de la expresión de $dszE$ en el fósrido completo. (Los detalles de la construcción se encuentran en M&M) pMPO1550, o en su defecto el plásmido vacío de la fusión, (pMPO234) se introdujeron en la estirpe MPO554Nal^R, previamente transformada con el fósrido UPO112. Estas estirpes se incubaron en medio mínimo con diferentes fuentes de azufre, durante 24 horas, tal y como se indica en materiales y métodos.

Como podemos observar en la figura 23, en ausencia de $dszE::lacZ$, se observan unos niveles espúreos de expresión que fluctúan en función de la fuente de azufre pero que en cualquier caso son bajos. Estos niveles son similarmente bajos aun en presencia de la fusión génica $dszE::lacZ$ cuando MPO554Nal^R / UPO112 / pMPO1550 crece en

presencia de sulfato, la fuente preferencial de azufre. Sin embargo, cuando crece a expensas de DBT, de taurina o de ambos, la expresión alcanza valores máximos, similares en las tres condiciones. Esto indica claramente que la expresión de *dszE*_{UPO112} está siendo regulada en función de la fuente disponible de azufre, asociando su expresión a fuentes de azufre limitantes, independientemente de la presencia de DBT, y reprimiéndose en presencia de una fuente preferencial. Un aspecto importante a considerar, es que, en un medio mínimo sin azufre donde la bacteria no puede crecer, la expresión es basal, sugiriendo que aún en condiciones limitantes, es necesario que haya crecimiento para que el operón *dszEABC* se exprese.

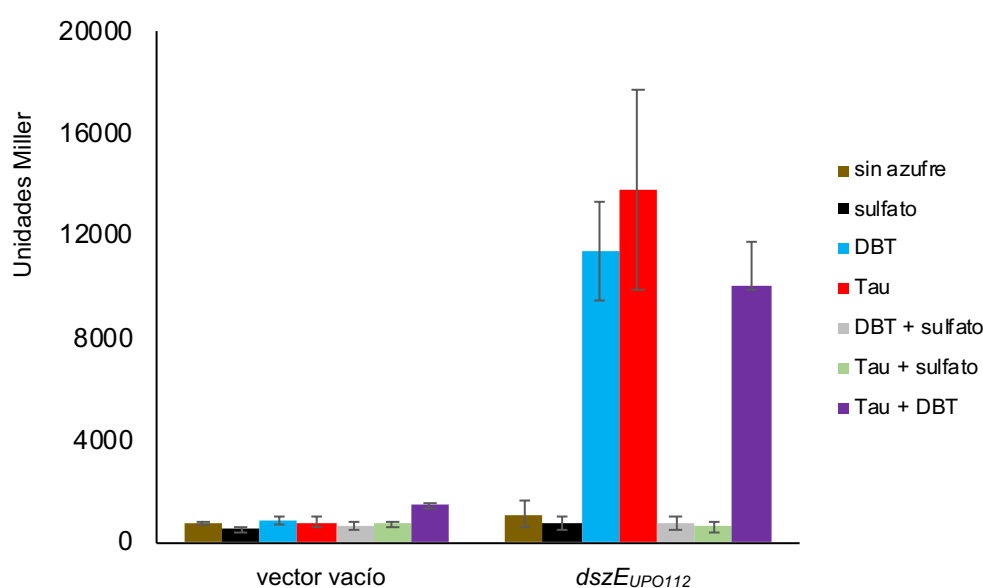


Figura 23: Ensayo de actividad betagalactosidasa de la fusión transcripcional *dszE*_{UPO112}::*lacZ* (pMPO1550), o el vector vacío de la fusión (pMPO234), en el fósido UPO112, en medio mínimo con diferentes fuentes de azufre. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

II.3.2. Transcripción de *dszE* y *dszR* dependiente de fuente de azufre.

Habiendo puesto de manifiesto la existencia de expresión regulada en función de la fuente de azufre, se quiso determinar si dicha regulación se produce exclusivamente a nivel de los propios genes catabólicos, o si la expresión del activador transcripcional de la ruta también estaba regulada. Para ello se realizaron experimentos de RT-PCR cuantitativa, amplificando fragmentos de *dszE* y *dszR*. Se emplearon muestras de cDNA obtenidas como se detalla en Materiales y Métodos a partir de

cultivos en fase exponencial media, de MPO554NaI^R / UPO112, crecidos en medio mínimo con sulfato, taurina o DBT, como fuentes de azufre.

Los resultados, que se muestran en la figura 24, indican que apenas hay transcripción detectable de *dszE* en las muestras procedentes de medio mínimo con sulfato. La transcripción aumenta en presencia de taurina y aún mas cuando crece a expensas de DBT, alcanzándose unos valores máximos de aproximadamente 100 veces en presencia de DBT, y de 50 veces en presencia de taurina. Estos datos son plenamente coincidentes con lo observado utilizando la fusión génica *dszE::lacZ*, lo que confirma la represión por fuente rica de azufre. En el caso de *dszR*, los valores absolutos son mucho menores, lo que tiene sentido al tratarse de un gen regulador, pero de nuevo observamos máximos valores en taurina o DBT, los cuales son aproximadamente 5 veces mayores en ambos, con respecto a los obtenidos en sulfato.

Por tanto; la transcripción tanto de *dszE* como de *dszR* se asocia a fuentes limitantes de azufre.

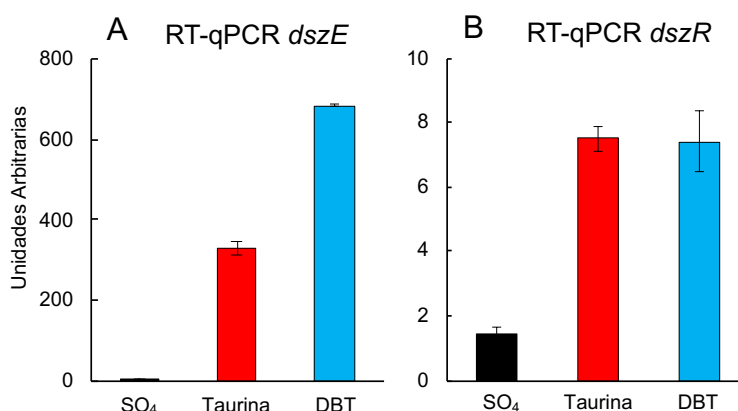


Figura 24: Ensayo de RT-qPCR con cDNA de cultivos de MPO554NaI^R / UPO112, crecidos en medio mínimo con diferentes fuentes de azufre. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos. **A:** niveles de transcripción de *dszE* cuando MPO554NaI^R / UPO112 crece en distintas fuentes de azufre. **B:** niveles de transcripción de *dszR* cuando MPO554NaI^R / UPO112 crece en distintas fuentes de azufre.

II.4. Papel de *dszHR* en la regulación por fuente de azufre.

II.4.1. Subclonación de *dszH-dszR_{wt}*.

Para profundizar más en el papel de DszR_{UPO112} como activador de la transcripción, se decidió subclonar *dszR_{UPO112}* en el mismo fósido pMPO579, y estudiar su expresión desde su propio promotor (sin arabinosa ni salicilato), para confirmar que *dszR_{UPO112}*, por sí mismo, es suficiente para activar la transcripción de los genes estructurales.

Puesto que aguas arriba de *dszR*, solapando con su ATG se encuentra *dszH*, y que muy probablemente se encuentren acoplados traduccionalmente, se decidió subclonar ambos genes, incluyendo 225pb aguas arriba del ATG de *dszH*, para tratar de incluir el promotor de dicho gen y establecer si *dszH* también estaba implicado en la expresión de *dszEABC*.

II.4.2. Regulación de *dszE::lacZ* por *dszHR_{wt}* en limitación de azufre.

En primer lugar se decidió medir la actividad betagalactosidasa producida por el plásmido de la fusión transcripcional *dszE_{UP0112}::lacZ* (pMPO1550) en MPO554NaI^R, con el plásmido con los genes *dszHR* (pMPO1552), o el vector vacío (pMPO579), en cultivos crecidos de igual manera que en el apartado II.3.2 de resultados, aunque en este caso no se estudió la expresión en DBT como fuente de azufre, puesto que el subclón no puede crecer en estas condiciones y, como se ha comentado, es necesario crecimiento para que haya expresión de *dszE*. Por tanto, sólo se estudió su expresión en sulfato, taurina o ambos simultáneamente, como fuentes de azufre.

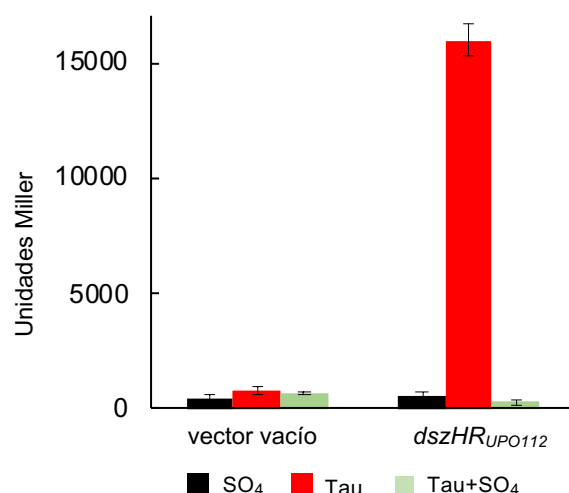


Figura 25: Ensayo de actividad betagalactosidasa de la fusión transcripcional *dszE_{UP0112}::lacZ* (pMPO1550), con el plásmido con los genes *dszHR* (pMPO1552), o el vector vacío (pMPO579), en medio mínimo con diferentes fuentes de azufre. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

Los resultados obtenidos se muestran en la figura 25, en la cual podemos observar solo niveles basales de expresión en cualquiera de las fuentes de azufre, en ausencia de *dszHR* (vector vacío pMPO579). Sin embargo, en su presencia (pMPO1552) se obtuvo un nivel muy alto de expresión (en torno a 16000 U. Miller) cuando la estirpe creció en la fuente limitante de azufre taurina, mostrando así la

capacidad de inducción por *dszHR*. En presencia de sulfato, se obtuvieron niveles basales de expresión, independientemente de que en el medio también esté presente taurina. Estos resultados, que son plenamente coincidentes con los obtenidos con el fósido completo, permiten deducir, que el fragmento subclonado en pMPO1552 incluye el promotor de *dszHR* y confirman que la expresión de *dszE* es dependiente de *dszHR*, y que dicha expresión está reprimida por fuente rica de azufre.

II.4.3. Expresión de *dszEABC* en limitación de nitrógeno.

Hasta el momento, se ha asociado la expresión de los genes *dsz* a condiciones de limitación de azufre. Pero, para asegurar que esta afirmación es cierta, y que la expresión no es debida a un déficit general en el estado nutricional, se decidió comprobar si los genes *dsz* podían expresarse también en otras condiciones limitantes de crecimiento diferentes a la fuente de azufre, por ejemplo, la limitación de nitrógeno.

Para ello se creció la estirpe MPO554NaI^R con los plásmidos pMPO1550 y pMPO1552 en medio mínimo con amonio como fuente rica de nitrógeno, y con glutamina o arginina como fuentes limitantes, además de sulfato o taurina, como fuentes de azufre, en cada caso. Los resultados se muestran en la figura 26.

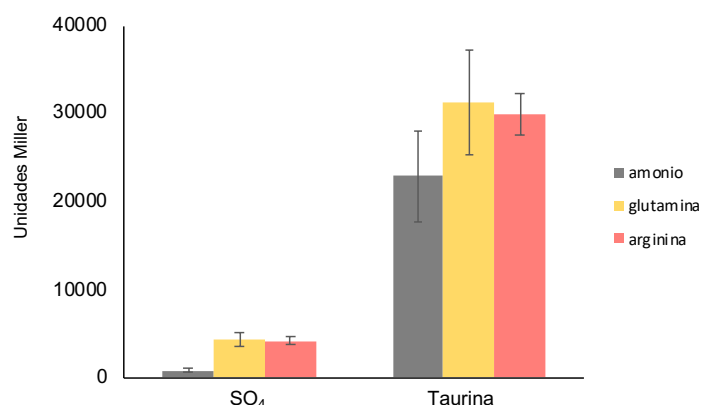


Figura 26: Ensayo de actividad betagalactosidasa de la fusión transcripcional *dszEUP0112::lacZ* (pMPO1550), con el plásmido con los genes *dszHR* (pMPO1552), en medio mínimo con diferentes fuentes de nitrógeno y de azufre. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos

Tal y como se aprecia en la figura, aunque en medio mínimo con sulfato se observa cierta diferencia en la expresión en función de la fuente de nitrógeno, (aproximadamente 5 veces más en glutamina o arginina con respecto a amonio), en el caso de la limitación de azufre, los niveles de expresión son mucho mayores, y apenas se observa efecto de la limitación de nitrógeno (1,3 veces más de expresión en glutamina o arginina con respecto a amonio), con lo que podemos concluir que si hay

un efecto debido al déficit general en el estado nutricional, este es mínimo, con respecto al efecto de la limitación de azufre.

II.4.4. Delección interna en *dszH*.

Para tratar de establecer si *dszH* tiene un papel en la activación de la expresión de *dszEABC* en limitación de azufre, se construyó un mutante de delección interna en fase sobre pMPO1552 (obteniendo pMPO1561). La delección elimina 188 codones de la parte central de *dszH*, conservando en pMPO1561 45 pb del extremo 5' y 42 pb del extremo 3' del gen. La extensión de la delección generada fue confirmada por secuenciación.

pMPO1561 se introdujo en MPO554NaI^R previamente transformada con pMPO1550 y se estudió su expresión en condiciones de limitación de azufre, de igual forma que se había hecho previamente para pMPO1552.

Los resultados se muestran en la figura 27 en la que observamos que, en ausencia de inductores, la activación se ve completamente abolida en presencia de la delección de *dszH*, independientemente de la fuente de azufre, puesto que los niveles de expresión son similares a los del plásmido vacío. Lo cual, sugiere una total dependencia de *dszH*, para que se produzca la activación.

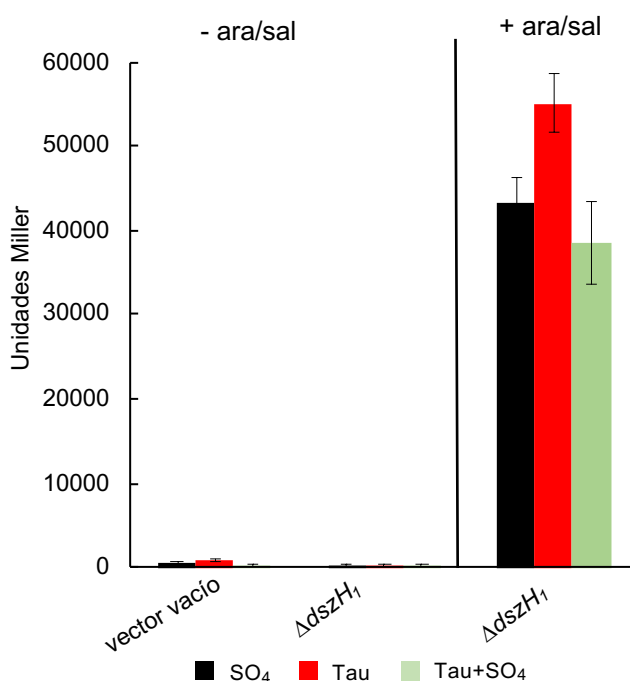


Figura 27: Ensayo de actividad betagalactosidasa de la fusión transcripcional *dszE_{UPO112}::lacZ* (pMPO1550), con el plásmido con la delección interna en fase en *dszH* ($\Delta dszH_1$, pMPO1561), en medio mínimo con diferentes fuentes de azufre. En el panel izquierdo se muestran los valores sin aporte de inductores, y en el panel derecho los valores con adición de arabinosa y salicilato.

Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

Sin embargo, al sobreexpresar *dszR* desde el promotor heterólogo (con arabinosa y salicilato) figura 27 (panel derecho), observamos altos niveles de inducción, sin discriminación de la fuente de azufre, lo que podría indicar que una mayor cantidad de DszR podría suplir la falta de DszH.

Por otra parte, al estudiar por RT-qPCR la transcripción de *dszR* en el plásmido con la versión silvestre de *dszHR* (pMPO1552) vs. el plásmido con la delección interna en *dszH* (pMPO1561), de igual manera que se había realizado para el fósido UPO112 se observa que solo hay una pequeña diferencia entre los niveles de transcripción de *dszR* en un plásmido y el otro (figura 28), la cual por sí misma no parece ser responsable de la falta de activación de la transcripción de *dszE*, cuando *dszH* está delecionado. Por tanto, el efecto de DszH sobre la función activadora de DszR, debe ser en un paso posterior a la transcripción de DszR.

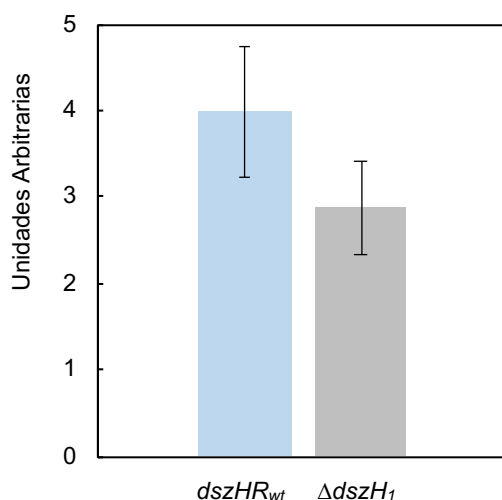


Figura 28: Niveles de transcripción de *dszR*, determinados por RT-qPCR con cDNA de cultivos de MPO554NaI^R / *dszHR_{wt}* (pMPO1552) y de MPO554NaI^R / $\Delta dszH_1$ (pMPO1561), crecidos en medio mínimo con taurina como fuente de azufre. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

A su vez, en un intento de esclarecer la posible función de DszH, se ha tratado de conseguir inducción de *dszE* intentando complementar la delección interna de *dszH* con *dszH* proporcionado en *trans*. Para ello se construyó un plásmido que lleva clonado *dszH* sin su propio promotor, pero que se transcribe bajo el control de P_{tac} (pMPO1571). Como se muestra en la figura 29, la función de DszH no pudo complementarse en *trans*,

lo que sugiere que no es la falta de la proteína DszH en sí misma lo que impide la expresión de *dszE*.

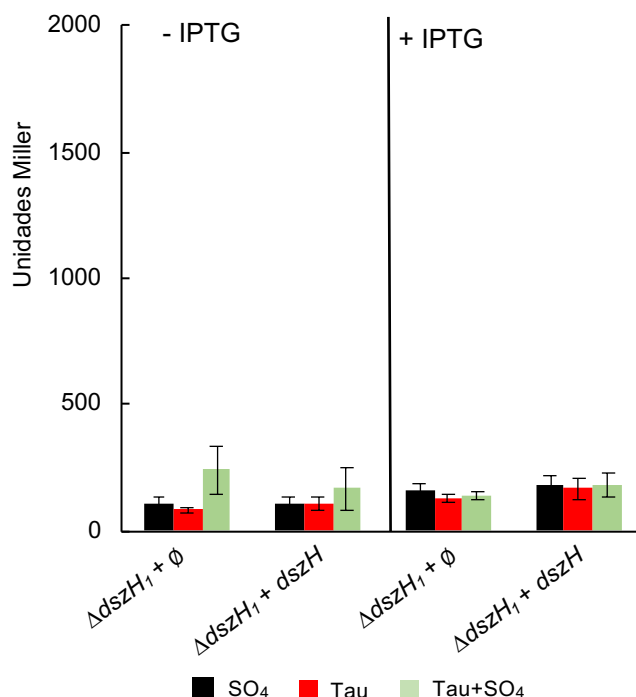


Figura 29: Ensayo de actividad betagalactosidasa de la fusión transcripcional *dszE*_{UPO112::lacZ} (pMPO1550), con el plásmido con la delección interna en fase en *dszH* ($\Delta dszH_1$, pMPO1561) y la versión de *dszH* sin su promotor (bajo P_{tac}) (pMPO1571), o en su defecto el plásmido vacío (\emptyset , pIZ1016), en medio mínimo con diferentes fuentes de azufre. En el panel izquierdo se muestran los valores sin inducir con IPTG, y en el panel derecho los valores con adición de IPTG. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

II.4.5. Expresión de *dszR* desde promotores heterólogos.

Hasta el momento no hemos podido discriminar si la regulación por azufre del operón estructural se debe al control transcripcional del operón *dszHR* o a una respuesta directa de la actividad reguladora de DszR.

Para tratar de aclarar estos interrogantes, se construyeron versiones mutantes carentes del promotor del operón *dszHR*, en los que el operón se transcribía bajo el control de P_{sal} (plásmido pMPO1560), y bajo el control de P_{tac} (pMPO1559). Por otra parte, se construyeron plásmidos carentes de P_{dszHR} que además tenían una delección interna en fase en *dszH*, en este caso algo menor que la de pMPO1561 (plásmido pMPO1564, en el caso de P_{sal} y pMPO1563 en el caso de P_{tac}), en la que se eliminan 176 codones de la parte central *dszH*, conservando 60 pb del extremo 5', y 63 pb del extremo 3'. Esta delección se hizo más pequeña, por si el drástico efecto observado en

la expresión de pMPO1561 era debido a una aberración en la construcción que afectara la expresión de DszR, más que a un efecto regulador.

En la figura 30 se muestra la expresión de *dszEUPO₁₁₂* cuando ambas versiones del operón regulador (ΔP_{dszHR} y $\Delta P_{dszHR}\Delta dszH_2$) se expresan exclusivamente desde P_{sal} .

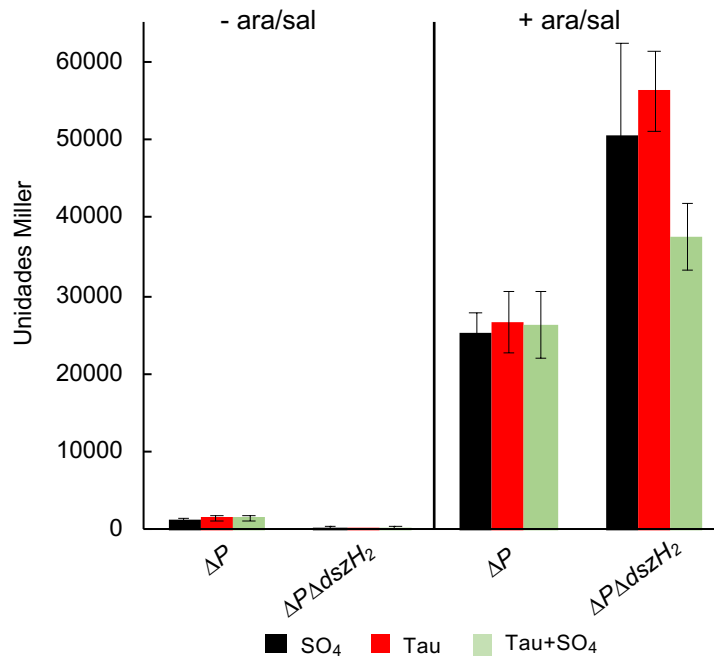


Figura 30: Ensayo de actividad betagalactosidasa de la fusión transcripcional *dszE_{UPO112}::lacZ* (pMPO1550), con el plásmido con la delección del promotor de *dszHR* (ΔP_{dszHR} , pMPO1560), y con el plásmido con la delección interna en fase en *dszH* pequeña y con el promotor de *dszHR* deleccionado ($\Delta P_{dszHR} \Delta dszH_2$, pMPO1564); ambos bajo el control de P_{sal} . En medio mínimo con diferentes fuentes de azufre. En el panel izquierdo se muestran los valores sin inductores, y en el panel derecho los valores con adición de arabinosa y salicilato. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

En ausencia de inductores, es decir, en baja copia y sin salicilato, no es posible la inducción de *dszE* en ninguna fuente de azufre. Al producir DszR de forma heteróloga, induciendo con arabinosa y salicilato (panel derecho) observamos altos niveles de expresión en la versión ΔP_{dszHR} , que son independientes de la fuente de azufre, y aún mayores (e igualmente independientes de fuente de azufre) cuando ambas delecciones están presentes. Estos datos sugieren por un lado que cuando DszR se produce en gran cantidad desde otro promotor distinto al suyo, la expresión de *dszE* ya no está regulada por azufre y, por otro, que la función de DszH sería prescindible.

Cuando las versiones del operón regulador se expresan desde P_{tac} , (figura 31), se observa que el nivel basal de este promotor (que es muy superior al de P_{sal}), sin necesidad de añadir IPTG, parece ser suficiente como para permitir la transcripción de *dszE* con niveles bastante altos e independientes de la fuente de azufre, siempre que *dszHR* estuviera completo. Sin embargo, con la delección interna de *dszH*, la transcripción de *dszE* se redujo drásticamente en ausencia de IPTG. Estos datos refuerzan por una parte la idea de que *dszH* debe ser necesario para la expresión de *dszEABC* y, por otra, sugieren que, si la expresión del operón *dszHR* no está regulada por azufre, la expresión de *dszE* tampoco lo está, aunque *dszH* esté presente.

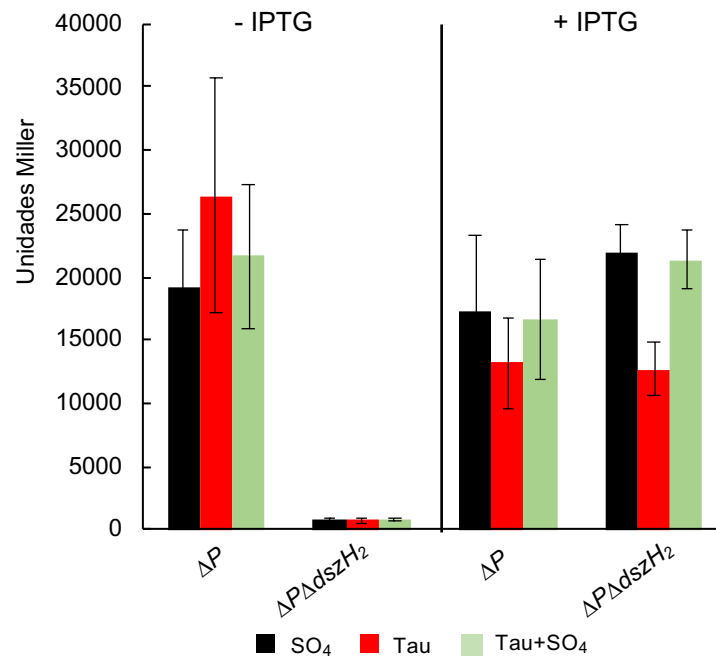


Figura 31: Ensayo de actividad betagalactosidasa de la fusión transcripcional *dszE_{UPO112}::lacZ* (pMPO1550), con el plásmido con la delección del promotor de *dszHR* (ΔP_{dszHR} , pMPO1559), y con el plásmido con la delección interna en fase en *dszH* pequeña y con el promotor de *dszHR* delecionado ($\Delta P_{dszHR} \Delta dszH_2$, pMPO1563); ambos bajo el control de P_{tac} , en medio mínimo con diferentes fuentes de azufre. En el panel izquierdo se muestran los valores sin inducir, y en el panel derecho los valores con adición de IPTG. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

Por otra parte, al inducir ambos plásmidos con IPTG (panel derecho), observamos altos niveles de expresión, similares en las tres fuentes de azufre, al igual que sucedía al sobreexpresar *dszHR* bajo el control de P_{sal} , apoyando de nuevo la idea, de que altos niveles de DszR, pueden soslayar la necesidad de DszH, y confirmando

que DszR puede activar la transcripción de *dszEABC* independientemente de la fuente de azufre si se transcribe ectópicamente.

III. Construcción de casetes de biodesulfuración de DBT de amplio espectro.

Puesto que este compuesto supone un grave problema de contaminación en los combustibles fósiles que ha de ser eliminado durante el proceso de refinado antes de su combustión, una vez caracterizados los genes catabólicos y reguladores de DBT y observado que parecen funcionar bien en un contexto genético distinto al propio, se decidió construir distintos vectores de amplio espectro capaces de expresar los genes de DBT en diferentes estirpes, con objeto de aunar en una misma estirpe bacteriana la capacidad degradadora de DBT con otras características que pudieran tener estas estirpes, como formación de biofilm o resistencia a solventes orgánicos, que permitiesen diseñar biocatalizadores degradadores de DBT con mayor potencialidad de ser utilizados en un proceso industrial.

III.1. Toxicidad de 2-HBF en *E. coli* y en *Pseudomonas*.

En primer lugar, es necesario tener en cuenta que en la vía 4S de degradación de DBT se produce la liberación de 2-hidroxibifenilo, el cual es un bactericida de características similares al triclosán (Czechowska, *et al.*, 2013). Además, se ha observado que la acumulación de 2-HBF puede dificultar el crecimiento y degradación de DBT por parte de bacterias degradadoras aisladas de distintas fuentes (Abin-Fuentes *et al.*, 2013); por tanto, para elegir estirpes en las que testar los casetes catabólicos que se pretendían construir, era necesario estudiar previamente la tolerancia de las diferentes estirpes a dicho compuesto.

Para ello se emplearon *E. coli* MPO554NaI^R, estirpe de referencia que se ha venido utilizando en todo este trabajo, así como 2 estirpes del género *Pseudomonas*, habitualmente empleadas como organismos modelo en estudios de degradación de compuestos aromáticos, la estirpe KT2440 y la estirpe DOT-T1E, que tiene una alta tolerancia a solventes orgánicos y puede ser empleada en cultivos bifásicos, (Rojas *et al.*, 2004), lo cual es de gran interés para aumentar la eficiencia del proceso de biodesulfuración, (Kawaguchi, *et al.* 2012).

Los estudios de toxicidad se realizaron según se detalla en materiales y métodos. En el caso de *E. coli* se realizaron tanto en medio líquido como en medio sólido, y en el caso de *Pseudomonas*, se realizaron exclusivamente en medio sólido.

III.1.1. Toxicidad en medio líquido.

Como se ha comentado, además del propio efecto tóxico del 2-HBF para la bacteria, su acumulación es un factor limitante en los procesos de BDS, causando inhibición por producto, por eso, se testó también el efecto del 2-HBF, en MPO554NaI^R

transformada con UPO112, para poder observar si la adición inicial de 2-HBF limitaba particularmente el crecimiento de dicha estirpe utilizando DBT como única fuente de azufre o el efecto tóxico sería general independientemente de la fuente de azufre.

Se realizaron curvas de crecimiento de MPO554NaI^R / UPO112 en medio mínimo con sulfato, taurina o DBT como fuentes de azufre, añadiendo además en cada fuente de azufre 0,2 mM de 2-HBF, (cantidad máxima teórica de liberación de 2-HBF a partir de 0,2 mM de DBT) y 0,4 mM de 2-HBF.

Los resultados, representados en la figura 32, muestran que con 0,2 mM de 2-HBF presente en el medio el crecimiento es casi igual al que se produce en ausencia de 2-HBF, con un ligero efecto quizá en las fases finales en las que los cultivos no alcanzan el máximo de densidad óptica. Este ligero efecto se observa en todos los cultivos, independientemente de la fuente de azufre. Al incrementar la concentración de 2-HBF hasta 0,4 mM, si que se observó una clara inhibición del crecimiento, cuantificándose en torno a un 30% cuando utiliza taurina o sulfato como fuentes de azufre, y algo más, en torno al 40 % cuando utiliza DBT.

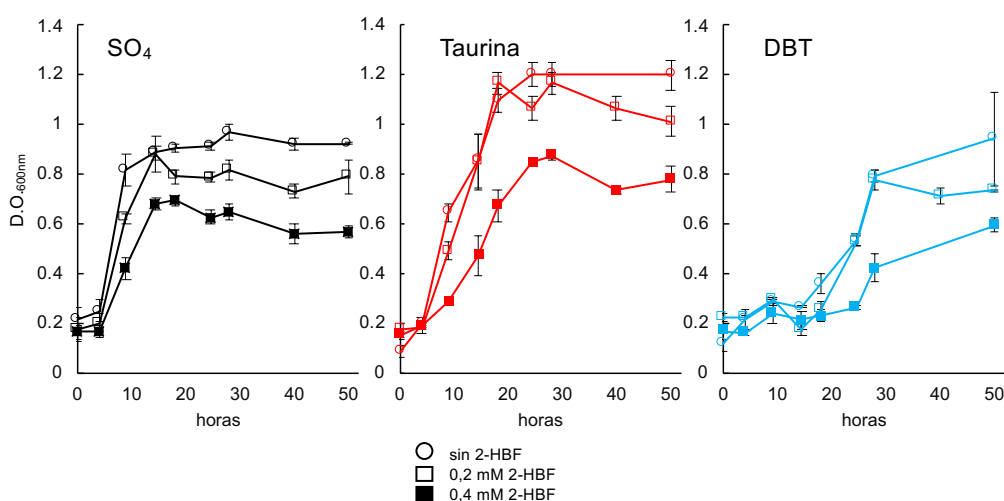


Figura 32: Curvas de crecimiento de MPO554 NaI^R / UPO112 en medio mínimo con diferentes fuentes de azufre, y con adición de 0,2 mM (cuadrados vacíos), 0,4 mM de 2-HBF (cuadrados rellenos) o sin adicionar 2-HBF (círculos). Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

III.1.2. Toxicidad en medio sólido.

Se realizaron a su vez ensayos de supervivencia con diluciones seriadas y siembra en gotas, tal y como se describe en materiales y métodos, con las estirpes *E. coli* MPO554NaI^R y de *P. putida* KT2440 y DOT-T1E en sus medios mínimos correspondientes, con sulfato como fuente de azufre y diferentes concentraciones de 2-HBF presente. Aunque uno de los casetes se probó también en la estirpe *P. putida* MPO406, que es superproductora de biofilm, característica de interés para mejorar el

proceso de BDS (Martínez *et al.*, 2017), en este caso no se realizó estudio de toxicidad previo, dada la dificultad de realizar el ensayo de supervivencia en gotas en una estirpe formadora de biofilm.

Tal y como se observa en la figura 33, con la adición de 0,2 mM de 2-HBF no se observó reducción de la viabilidad en ninguna de las tres estirpes (datos no mostrados). *E. coli* presentó cierta sensibilidad a la adición de 0,4 mM de 2-HBF y una sensibilidad total con la adición de 0,6 mM. Las estirpes de *Pseudomonas*, sin embargo, son bastante más tolerantes al tóxico, no mostrando reducción alguna de la viabilidad incluso a la máxima concentración de 2-HBF.

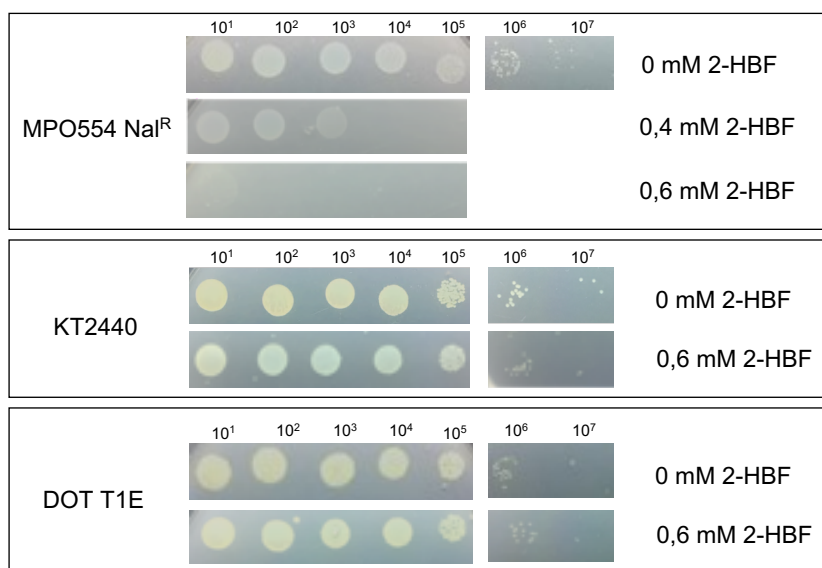


Figura 33: Resultado primario de un experimento representativo de un ensayo de supervivencia con diluciones seriadas y siembra en gotas, en medio mínimo con sulfato como fuente de azufre y con adición de diferentes concentraciones de 2-HBF.

III.2. Diseño y construcción de casetes catabólicos de amplio espectro.

Para desarrollar casetes catabólicos de uso industrial, se necesitaba contar con vectores de amplio espectro que permitieran la expresión del sistema en distintas estirpes. Además, se quisieron construir vectores en los que el operón *dszEABC* se expresara por sí mismo, utilizando su propio sistema regulador, pero también otro en el que la expresión del operón pudiera regularse por medio de un promotor distinto. Para ello se contó con los vectores de amplio espectro del sistema SEVA: pSEVA221 y pSEVA429 (Silva-Rocha *et al.*, 2013), que cuentan con *oriV*, *trfA* y *oriT*, por lo que pueden ser transferidos por conjugación y mantenidos en distintas bacterias Gram negativas. En el caso de pSEVA429, además cuenta con el sistema de expresión *alkS*-*P_{alkB}*, inducible por alcanos. Los plásmidos se construyeron a partir de UPO112 tal y

como se detalla en M&M y se esquematizan en la figura 34. El plásmido cuya expresión depende de su propio sistema (pMPO1566), contiene a su vez *dszHR* con su propio promotor, en la misma orientación que los genes estructurales, y separados por 1082 pb. El plásmido bajo el control de *alkS*- P_{alkB} (pMPO1570), contiene los genes estructurales sin su propio promotor.

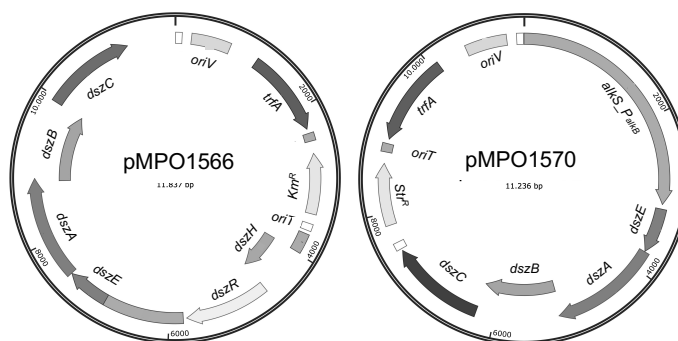


Figura 34: Mapa genético de los plásmidos pMPO1566 y pMPO1570, donde se indican los genes de UPO112 clonados en los vectores originales pSEVA221 y pSEAV429, respectivamente, así como los elementos originales de los plásmidos.

III.3. Validación en *E. coli* y *Pseudomonas*.

Para validar el funcionamiento de los casetes construidos, los plásmidos pMPO1566 y pMPO1570 (o sus correspondientes vectores vacíos) se conjugaron a la estirpe *E. coli* MPO554Nal^R y a las estirpes de *P. putida* KT2440 y DOT-T1E, en el caso de pMPO1566, y a las estirpes de *P. putida* KT2440, DOT-T1E y MPO406, en el caso de pMPO1570. Las estirpes resultantes se crecieron en sus correspondientes medios mínimos con DBT como única fuente de azufre, llevando además como control un medio sin azufre añadido, para poder asignar el crecimiento, en caso de que se produjera, a la degradación de DBT. En este punto y en el caso de *Pseudomonas*, fue necesaria la utilización de un medio mínimo muy pobre, y una pre-incubación sin azufre, tal y como se detalla en M&M, para asegurar el total agotamiento del azufre acumulado por la bacteria durante el preinóculo. Para confirmar, además, que la degradación de DBT se estuviera produciendo por la vía 4S, se determinó el 2-HBF producido en el medio por el método de Gibbs, según se detalla en M&M.

III.3.1. Expresión desde su propio promotor.

En primer lugar, se procedió a validar el plásmido pMPO1566, que expresa los genes de degradación de DBT a partir de sus propias señales, y debe ser por tanto dependiente de DBT.

El plásmido se introdujo en las estirpes: *E. coli* MPO554NaI^R, *P. putida* KT2440 y *P. putida* DOT-T1E, los resultados del crecimiento de las estirpes resultantes en medio mínimo con DBT se muestran en la figura 35, y los de la producción de 2-HBF en la tabla 11.

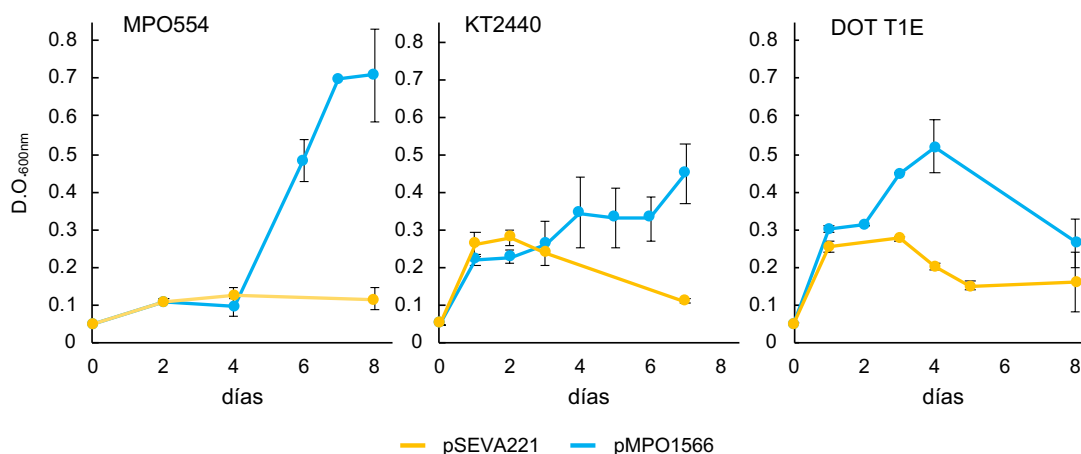


Figura 35: Curvas de crecimiento de MPO554 NaI^R, KT2440 y DOT T1E con pMPO1566 o el vector vacío pSEVA221, en medio mínimo con DBT como única fuente de azufre. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

A pesar de la incubación previa sin azufre, las estirpes de *Pseudomonas* con el plásmido vacío o con pMPO1566 en medio mínimo sin azufre (datos no mostrados), presentaron cierto crecimiento residual.

MPO554NaI^R / pMPO1566, presentó una fase lag de 4 días, pero posteriormente creció, aunque más lentamente que con el fósido UPO112 original, alcanzando finalmente una DO₆₀₀ de sólo 0,7 al cabo de los 7 días, mientras que con el fósido completo UPO112 se alcanza el máximo, a una DO₆₀₀ de 1, en 48 h (figura 3). Al final de la incubación el cultivo había acumulado 0,053 mM de 2-HBF (tabla 11), lo que sugiere que no había consumido todo el DBT.

La estirpe silvestre *P. putida* KT2440 presentó un crecimiento lento pero continuo, alcanzando tras 7 días una DO₆₀₀ de 0,45 (figura 3), con una producción de 2-HBF de 0,101 mM (tabla 11).

En cuanto a la estirpe *P. putida* DOT-T1E, alcanzó una DO₆₀₀ de 0,52, en sólo 4 días, aunque a partir de ese tiempo la DO₆₀₀ incluso decreció (figura 3). Su producción final de 2-HBF fue la mayor de las tres (0,139 mM) (tabla 11).

| Estirpe | Fuente de azufre | 2-HBF (mM) |
|------------------|---------------------|-------------|
| MPO554/pSEVA221 | DBT | 0,000±0,000 |
| MPO554/pMPO1566 | DBT | 0,053±0,006 |
| KT2440/pSEVA221 | DBT | 0,000±0,000 |
| KT2440/pMPO1566 | DBT | 0,101±0,005 |
| KT2440/pMPO1566 | DBT+SO ₄ | 0,004±0,001 |
| DOT T1E/pSEVA221 | DBT | 0,001±0,000 |
| DOT T1E/pMPO1566 | DBT | 0,139±0,026 |
| DOT T1E/pMPO1566 | DBT+SO ₄ | 0,004±0,003 |

Tabla 11: Producción de 2-HBF con pMPO1566 o pSEVA221 (medio mínimo con DBT o con DBT y sulfato).

En su conjunto, estos datos indican que el operón *dszEABC* ha podido expresarse desde sus propias señales en las estirpes de *Pseudomonas*, al igual que en *E. coli*.

Con objeto de establecer si la expresión desde sus propias señales está también sujeta a represión por sulfato en *Pseudomonas*, las dos estirpes de este género se crecieron también en medio mínimo con DBT y sulfato. En estas condiciones el crecimiento fue superior, alcanzando el cultivo su máxima densidad (en torno a 1-1,2), al cabo de 24 horas (no mostrado). La cantidad de 2-HBF acumulada por estos cultivos, que se midió tras el mismo tiempo de incubación que la de los cultivos que sólo contenían DBT, fue prácticamente indetectable, tal y como se observa en la tabla 11, lo que sugiere que la capacidad degradadora de DBT, se reprime por SO₄ en *Pseudomonas*, al igual que en *E. coli*.

III.3.2. Expresión desde el sistema *alkS*-*P_{alkB}*.

Una vez validado el plásmido pMPO1566, se testó pMPO1570 que expresa los genes de degradación de DBT a partir del promotor *P_{alkB}* regulado por *alkS*. En este caso, sólo se emplearon las estirpes de *Pseudomonas*, ya que *E. coli* MPO554NaI^R es también resistente al marcador del plásmido (estreptomicina).

Los datos de crecimiento se muestran en la figura 36 y los de producción de 2-HBF en la tabla 12.

Como se observa en la figura, tanto KT2440, como DOT T1E fueron capaces de crecer en sólo 3 días a expensas de DBT. Sorprendentemente, crecieron incluso los cultivos sin el inductor del sistema DCPK. De hecho, alcanzaron mayor DO₆₀₀ sin el inductor DCPK y, tal y como se muestra en la tabla 12, acumularon mayor cantidad de 2-HBF (1,3 veces más en el caso de KT y 1,5 veces más en el caso de DOT T1E).

Por el contrario, en el caso de MPO406 los cultivos sin inductor crecieron inicialmente de forma similar a los que tenían inductor añadido, pero dejaron de crecer a una DO = 0,35 de DO, mientras que la DO del cultivo con inductor continuó

aumentando hasta una DO = 0,62 al undécimo día. Esto contrasta con la producción de 2-HBF que en este punto fue de 0,129 en DBT sin DCPK y de 0,104 mM con DCPK (tabla 12). Puesto que no se ha realizado en paralelo un recuento de viables en placa, no se puede descartar que el aumento de la DO sea debido a la producción de componentes de la matriz extracelular, quizá en respuesta a la toxicidad del inductor, y no a un crecimiento real.

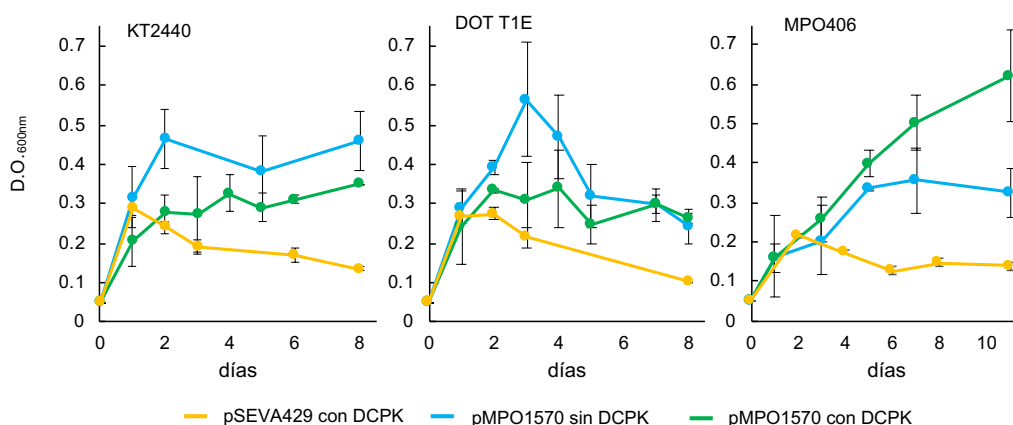


Figura 36: Curvas de crecimiento de KT2440, DOT T1E y MPO406 con pMPO1570 o el vector vacío pSEVA429, en medio mínimo con DBT como única fuente de azufre. Los valores representados son la media de 3 réplicas. Las barras de error representan la desviación estándar de los datos.

| Estirpe | DCPK | Fuente de azufre | 2-HBF (mM) |
|------------------|------|----------------------|-------------|
| KT2440/pSEVA429 | - | DBT | No testado |
| KT2440/pSEVA429 | + | DBT | 0,000±0,000 |
| KT2440/pMPO1570 | - | DBT | 0,114±0,027 |
| KT2440/pMPO1570 | + | DBT | 0,082±0,008 |
| KT2440/pMPO1570 | - | DBT+ SO ₄ | 0,036±0,002 |
| KT2440/pMPO1570 | + | DBT+ SO ₄ | 0,014±0,002 |
| DOT T1E/pSEVA429 | - | DBT | 0,002±0,001 |
| DOT T1E/pSEVA429 | + | DBT | 0,001±0,000 |
| DOT T1E/pMPO1570 | - | DBT | 0,133±0,030 |
| DOT T1E/pMPO1570 | + | DBT | 0,088±0,017 |
| DOT T1E/pMPO1570 | - | DBT+ SO ₄ | 0,036±0,006 |
| DOT T1E/pMPO1570 | + | DBT+ SO ₄ | 0,014±0,003 |
| MPO406/pSEVA429 | - | DBT | No testado |
| MPO406/pSEVA429 | + | DBT | 0,000±0,000 |
| MPO406/pMPO1570 | - | DBT | 0,129±0,034 |
| MPO406/pMPO1570 | + | DBT | 0,104±0,024 |
| MPO406/pMPO1570 | - | DBT+ SO ₄ | 0,042±0,004 |
| MPO406/pMPO1570 | + | DBT+ SO ₄ | 0,042±0,005 |

Tabla 12: Producción de 2-HBF con pMPO1570 o pSEVA429 (medio mínimo con DBT o con DBT y sulfato).

Está descrito que este sistema de regulación tiene un “considerable” nivel basal, (Arce-Rodríguez *et al.*, 2020), y parece ser que, en este caso, con el número de copias del vector pSEVA, este nivel basal de transcripción de *dszEABC* desde P_{alkB} es suficiente como para que la capacidad de biodesulfuración de DBT no sea el factor limitante del crecimiento a expensas de DBT.

Por otra parte, si consideramos los datos de producción de 2-HBF en presencia de DBT vs DBT y sulfato (tabla 12), que se midió tras el mismo tiempo de incubación en ambas condiciones (el necesario para que los cultivos con DBT como única fuente de azufre alcanzaran fase estacionaria), podemos observar que en presencia de sulfato y DBT se produce una cantidad de 2-HBF aproximadamente 3 veces menor, en comparación con el mismo cultivo con DBT como única fuente de azufre, en las tres estirpes. Esta reducción podría deberse al rápido crecimiento de las bacterias en estas condiciones, que no daría tiempo a una completa eliminación del DBT antes de llegar a fase estacionaria, puesto que alcanzaron su máxima DO (en torno a 1,2) en 24h (datos no mostrados). En cualquier caso, el hecho de que con este sistema de expresión ectópico se redujese la producción de 2-HBF solo 3 veces en presencia de sulfato, mientras que con su propio sistema de expresión la presencia de sulfato elimine completamente su producción, apoya la idea de que su sistema de expresión está también reprimido en *Pseudomonas*, de forma similar a lo que ocurre en *E. coli*.

DISCUSIÓN

I. Características diferenciadoras de los genes *dsz* en UPO21 y UPO112.

Importancia de los estudios de metagenómica funcional.

Al realizar una búsqueda de clones con capacidad de crecer con DBT como única fuente de azufre, en una metagenoteca construida a partir de suelo contaminado procedente de una refinería de una industria petroquímica (Terrón-González *et al.*, 2016), se obtuvieron 20 candidatos capaces de crecer en menos de 48 horas en medio mínimo con DBT como única fuente de azufre (tabla 4). Todos los candidatos resultaron ser el mismo (UPO112), lo cual no significa que no existan otros clones con esta capacidad en la metagenoteca, pero si indica que la forma en que se realizó la búsqueda favoreció que se aislara probablemente el clon que presenta un crecimiento más rápido. UPO112 contiene en su inserto metagenómico los genes que codifican para las enzimas de la ruta de desulfuración de DBT, por la vía 4S: dibenzotiofeno-5,5-dióxido monooxigenasa (*dszA*), 2-hidroxibifenilo-2-sulfinato desulfinaasa (*dszB*) y dibenzotiofeno monooxigenasa (*dszC*) (Gray *et al.*, 1996) (tabla 5). Los genes *dsz* de UPO112 son similares a los previamente descritos en UPO21 (Terrón-González, 2014, Martín-Cabello *et al.*, 2020), el cual procede de la misma metagenoteca, pero en su caso, se aisló en una búsqueda de clones con actividad oxigenasa inicial y sólo mostró actividad productora de índigo en presencia de arabinosa y salicilato.

Los operones *dsz* de UPO21 y de UPO112, tienen acoplado el gen que codifica para la FMN reductasa de la ruta (*DszE*), que realizaría el papel que en otras bacterias realiza *DszD*, proporcionando los electrones necesarios a *DszA* y *DszC* (Gray *et al.*, 2003). Puesto que, hasta el momento, en ninguna bacteria degradadora de DBT por la vía 4S caracterizada se había encontrado la reductasa de la ruta formando parte del operón *dsz*, es posible que esta configuración génica no sea frecuente en la naturaleza. Sin embargo, la realización de los rastreos funcionales (tanto la búsqueda de clones con actividad oxigenasa inicial, como la de clones con capacidad de crecer con DBT como única fuente de azufre), podrían haber favorecido el aislamiento de este tipo de operones. Teniendo en cuenta los precedentes que indican que las oxidoreductasas de *E. coli* no parecen poder donar los electrones a las monooxigenasas implicadas en biodesulfuración del DBT y que, de hecho, se han utilizado estirpes de *E. coli* para identificar flavin reductasas capaces de realizar dicha función (Ishii *et al.*, 2000b; Furuya *et al.*, 2004; Furuya *et al.*, 2005), resulta improbable que los operones *dsz* que no tuvieran reductasas asociadas pudieran haber mostrado actividad oxigenasa o crecimiento en DBT en *E. coli*.

Es bastante sorprendente que en las bacterias caracterizadas hasta la fecha la reductasa no esté acoplada al resto del operón, ya que, en las rutas de degradación de contaminantes, lo normal es que el sistema de transporte de electrones esté acoplado al resto de genes implicados en la ruta. Probablemente esto indique que se trata de un sistema relativamente nuevo, en el que los elementos se han ido ensamblando de manera secuencial, y puesto que la actividad de la reductasa puede suplirse en algunos casos por otras de la propia bacteria, tal y como se ha detallado en la introducción, dicho gen aún no estaría incorporado a la vía en las rutas de DBT descritas con anterioridad.

El hecho de que los clones UPO21 y UPO112 presenten dicha reductasa acoplada puede ser una ventaja competitiva que les permita (sobre todo a UPO112) un crecimiento más eficiente. Por ello, este set de genes puede ser interesante para un posible uso de la BDS en distintas bacterias, ya que, hasta la fecha, si se ha querido controlar el aporte de electrones a las monooxigenasas, se ha tenido que hacer construyendo sistemas híbridos que incorporasen *dszD* al resto de genes *dsz*, puesto que su ausencia es uno de los cuellos de botella a superar en la ingeniería de la biodesulfuración de DBT a través de la vía 4S (Martínez *et al.*, 2017).

Un operón similar al de UPO21 y UPO112, (incluyendo *dszE*) ha sido localizado en el genoma recientemente publicado de la bacteria *Comamonadaceae bacterium*, betaproteobacteria del orden *Burkholderiales*, la cual ha sido aislada en cultivos de enriquecimiento de aguas subterráneas contaminadas con hidrocarburos, (Figueroa-Gonzalez *et al.*, 2020). Los genes *dszEABC* de *Comamonadaceae bacterium* presentan una identidad mayor del 92%, con los de UPO112, y en particular *DszE_{UPO112}* y la proteína NADPH-dependent FMN reductase de *Comamonadaceae bacterium* son prácticamente idénticas (figura 8). Ambas proteínas son muy parecidas a las alcanosulfonato monooxigenasas, aunque la NADPH-dependent FMN reductasa de *Comamonadaceae bacterium* presenta en su extremo amino una secuencia que no se parece a las alcanosulfonato monooxigenasas, ni a *DszE_{UPO112}*. Teniendo en cuenta que la reductasa de *Comamonadaceae bacterium* se ha definido solo por homología de secuencia, es posible que se trate de un error en la anotación identificando un codón de inicio erróneo, más que de una diferencia real. *DszE_{UPO21}* difiere algo más de las dos proteínas anteriores; de hecho, al realizar un estudio filogenético, las otras dos aparecen en la misma rama y *DszE_{UPO21}* aparece en otra rama diferente, aunque muy próximas (figura 9). En dicho estudio filogenético, se pone de manifiesto que *DszE*, está más próxima a las reductasas dependientes de NADPH, pertenecientes a la familia *SsuE*, que a otras reductasas de las rutas de DBT identificadas hasta el momento, lo cual es muy interesante, porque podría indicar que *DszE* se podría haber incorporado a la ruta desde otras rutas de degradación, para hacer más eficiente al operón *dsz*.

Un origen distinto de los genes catalíticos y del gen del transportador de electrones fue postulado de forma similar por Furuya *et al.*, 2005, que sostienen que probablemente los genes de desulfuración de DBT de la bacteria *M. phlei* WU-F1, y el gen de la flavina reductasa necesaria para la ruta de DBT en dicha bacteria, deben haber evolucionado de forma independiente.

Además del acoplamiento de *dszE* al operón *dszABC*, tanto UPO21 como UPO112 presentan otra característica muy importante que les diferencia del resto de genes de desulfuración de DBT descritos hasta la fecha, y es la presencia en las cercanías del gen codificante del activador transcripcional de la ruta, DszR (Terrón-González, 2014, Martín-Cabello *et al.*, 2020), que se encuentra separado de *dszC* \approx 3,4 kb en UPO21 y \approx 4,2 en UPO112, en sentido contrario al operón *dszEABC* (figura 7), siendo la primera vez que se describe un activador transcripcional ligado a una ruta de desulfuración de DBT. DszR, presenta una gran similitud con SfnR de *Pseudomonas fluorescens* (figura 10), siendo ambos activadores transcripcionales dependientes de sigma⁵⁴ y muy próximos filogenéticamente (figura 11).

El hecho de que DszE sea parecida a SsuE, y que las vías de desulfuración de DBT y de sulfonatos alifáticos, estén muy relacionadas, tal y como se ha detallado en la introducción, permite suponer, que DszE ha sido incorporada al operón *dszABC*, a partir de dichas vías para conseguir un operón más eficiente. Por otra parte, la similitud entre DszR y SfnR que, al igual que SsuE, también está implicado en la utilización de sulfonatos alifáticos, así como el modelo regulador que se propone a continuación y que es similar al descrito para la biodegradación de sulfonatos alifáticos en *Pseudomonas*, también permiten suponer que el operón *dszEABC* de UPO21 y UPO112 habría incorporado dicho gen regulador a partir de esas vías de degradación para conseguir una expresión más eficiente. La ausencia de *dszR* en el genoma de *Comamonadaceae bacterium*, indicaría que este elemento regulador se habría incorporado con posterioridad a *dszE*, lo cual es bastante común en el ensamblaje de rutas de degradación de contaminantes, ya que este proceso habitualmente ocurre de forma secuencial, ensamblándose en primer lugar los genes catabólicos y produciéndose posteriormente el perfeccionamiento de la regulación específica y su acoplamiento a la fisiología celular (Cases & de Lorenzo, 2001).

A pesar de la similitud entre ambos fósmidos, los cuales comparten los genes *dszEABC*, *dszR*, y la región génica que separa a ambos, UPO112 crece en DBT mucho más eficientemente que UPO21, y al contrario que éste, sin necesidad de expresión heteróloga (figura 3). Además, no produce índigo (figura 5), motivo por el que no fue

identificado en la misma búsqueda en la que se aisló UPO21; por tanto, los genes de UPO112 deben tener alguna característica diferente de los de UPO21. El hecho de que se hayan podido aislar ambos clones, pone de manifiesto la ventaja de realizar diferentes estudios de metagenómica funcional, empleando el mayor número posible de sistemas de búsqueda, para así aumentar las posibilidades de encontrar actividades de interés que tengan características distintas. De hecho, la realización de estudios de metagenómica funcional puede ser interesante para encontrar enzimas más eficientes, ya que, aunque las enzimas implicadas en la vía 4S son muy conocidas, presentan algunas limitaciones que reducen la eficiencia del proceso. Por ejemplo, DszC de *R. erythropolis* IGTS8 tiene una baja eficiencia catalítica y una alta sensibilidad tanto a 2-HBF como a HBFS, (Abin-Fuentes *et al.*, 2013) mientras que, a su vez, la reacción catalizada por DszB es la que limita la velocidad general del proceso de BDS (Martínez *et al.*, 2016), además de que el suministro de electrones por parte de DszD, también parece ser un paso limitante de la velocidad (Gray *et al.*, 2003).

Teniendo en cuenta que los operones *dsz* de UPO21 y UPO112 presentan dos características muy importantes y diferentes a otros descritos previamente, como son el hecho de que *dszE* esté acoplado a *dszABC*, y que se ha podido describir el activador transcripcional de la ruta (DszR), se pone de nuevo de manifiesto la utilidad de realizar estudios de metagenómica funcional en la búsqueda de actividades enzimáticas con potencial interés industrial.

II. Transferencia horizontal de los operones *dszEABC* de UPO21 y UPO112.

Los genes *dszABC* en ocasiones están asociados a elementos móviles (Ishii *et al.*, 2000a); además el grado de conservación de dichos genes entre diferentes bacterias es muy alto por lo que se piensa que en la naturaleza debe ser habitual que se transmitan de unas bacterias a otras mediante transferencia horizontal (Mohebbi & Ball, 2008). En el caso de UPO21 y UPO112, ambos comparten un fragmento de ADN en el que se encuentran todos los genes *dsz*, tanto catabólicos como reguladores. Este fragmento debe constituir una isla genómica puesto que sus secuencias flanqueantes son diferentes entre sí. Si tenemos en cuenta que *Comamonadaceae bacterium* comparte el operón *dszEABC*, presentando gran similitud de sus genes *dszEABC* con los de UPO112 (tabla 9), y con regiones intergénicas prácticamente idénticas (tabla 7), podemos imaginar que este operón *dsz* habría adquirido mayor complejidad en una etapa posterior, incorporado el gen regulador desde otras rutas de degradación, para generar la isla genómica completa observada en los fósidos. Las secuencias

flanqueantes de los genes *dsz* son diferentes entre UPO21 y UPO112, pero en ambos casos los genes de los dos fósquidos fuera de la isla genómica muestran gran similitud con genes de *Candidatus* Macondimonas diazotrophica (tabla 5 y Terrón-González, comunicación personal). Además, la secuencia y la organización génica del inserto de UPO112 es muy similar a una zona continua del genoma de dicha bacteria (tabla 10), salvo en la zona de la isla genómica que parece insertada, adyacente a la cual existen genes completos o fragmentos génicos que codifican para enzimas asociadas a ADN móvil. Ello sugiere que una bacteria con un genoma muy parecido al de *Candidatus* Macondimonas diazotrophica muy probablemente haya adquirido los genes *dsz* por transferencia horizontal, los cuales se habrían insertado como una isla genómica. En el caso de UPO21 las secuencias flanqueantes a la isla genómica indican que este evento de transferencia horizontal se habría repetido de forma independiente en una bacteria con un genoma también muy similar, pero en este caso insertándose la isla en una localización diferente. *Candidatus* Macondimonas diazotrophica es una Proteobacteria perteneciente a las Gammaproteobacteria; que ha sido aislada de sedimentos marinos contaminados con vertidos de petróleo por todo el planeta pudiendo llegar a ser dominante sobre otras bacterias, alcanzado el 30% de la población en dichas zonas, pero, sin embargo, está ausente en sedimentos prístinos o en agua de mar (Karthikeyan *et al.*, 2019). Al adquirir los genes *dsz*, *Candidatus* Macondimonas diazotrophica obtendría una ventaja competitiva. Esta bacteria, que además tiene genes de fijación de nitrógeno, lo cual es importante porque la biodegradación del petróleo a menudo está limitada por nitrógeno, podría así obtener azufre y nitrógeno fácilmente de los hidrocarburos, lo que le ayudaría a imponerse sobre otras poblaciones en este entorno con exceso de carbono respecto al resto de nutrientes.

III. Regulación de la expresión de *dszEABC*.

Hasta el momento no se ha descrito ningún sistema de regulación de los genes *dsz* que responda a DBT y los pocos datos disponibles sobre regulación de los genes *dsz* indican que es muy probable que su regulación sea un mecanismo de respuesta general de la bacteria al estrés y a la limitación de azufre, más que una respuesta a la presencia del sustrato metabolizable.

Los factores implicados en la regulación transcripcional de los genes *dsz* son a su vez muy poco conocidos; de hecho, solo se han descrito, en *Gordonia*, dos proteínas reguladoras que se unen a la región promotora de dichos genes. Una de ellas es un regulador de la familia TetR, propuesta como activador transcripcional (Murarka *et al.*, 2019a, Murarka & Srivastava, 2019b) y la otra el regulador global WhiB1, que actuaría

como represor, (Murarka *et al.*, 2020). DszR es la primera proteína reguladora de genes *dsz* caracterizada en profundidad, habiéndose demostrado tanto *in vivo* (figura 13) como *in vitro* (figura 22), que es necesaria para la activación transcripcional del operón *dszEABC*. Además, queda demostrado que es un activador de promotores dependientes de σ^{54} (figura 22) que, como muchos de ellos, necesita el concurso de IHF para una correcta activación de la transcripción, (Shingler, 2011). Por otra parte, sus sitios palindrómicos de unión (figura 21), se encuentran situados en la misma cara de la hélice lo que es común en este tipo de activadores (García-Mauriño *et al.*, 2013) para facilitar la oligomerización en forma de hexámeros, que debe ser la forma activa de la proteína (Gao *et al.*, 2020).

A pesar de que la secuencia de DszR indica que es un activador transcripcional, los ensayos de actividad betagalactosidasa de las fusiones a *lacZ* mostrados en la figura 13, sugirieron inicialmente que también podría ser necesario para la traducción de las proteínas Dsz, ya que cuando el operón se transcribía desde el promotor heterólogo P_{sal} no había traducción en ausencia de DszR. Dado que la función de los únicos dominios de la proteína DszR estaban bien caracterizados en este tipo de activadores transcripcionales, resultaba muy improbable, por no decir imposible, que DszR tuviese una función postranscripcional. La explicación más lógica era que solo los transcritos provenientes de su promotor nativo P_{dsz} , pero no los provenientes de P_{sal} , se podían traducir, y que, de hecho, la transcripción desde P_{sal} producía interferencia con la transcripción desde P_{dsz} de forma que cuando los dos promotores estaban activos, el nivel de expresión era incluso menor que cuando solo estaba activo P_{dsz} . Los resultados mostrados en la figura 16 al realizar una delección del promotor propio de forma que P_{sal} queda reposicionado en su lugar, indicaron claramente que en esta situación DszR es completamente dispensable, con lo que su papel en la traducción queda descartado.

La ausencia de traducción cuando los transcritos provienen del promotor P_{sal} localizado aguas arriba, se debe a la formación de dos estructuras secundarias en la más estable de las cuales, se encuentran ocultos tanto el codón de inicio de *dszE*, como la SD, lo que impide su traducción (figura 17). Un fenómeno similar se ha descrito en detalle en el promotor *glnK* de *E. coli*, asociando la falta de traducción a la presencia de secuencias cortas ricas en CT solapantes con las secuencias promotoras, que formarían estructuras secundarias silenciadoras de la traducción si la transcripción proviene de aguas arriba del propio promotor. (Levy *et al.*, 2017). Este mecanismo, que parece estar bastante extendido en promotores dependientes de σ^{54} (Levy *et al.*, 2017), supondría una excelente forma de aislar la expresión del operón, haciéndola completamente independiente del contexto genómico en el que se encuentre, lo cual sería una

característica de particular de interés en el caso de un operón móvil que se transfiere horizontalmente y que se inserta en diferentes genomas. El sistema de aislamiento sería completamente eficaz si todos los genes del mismo operón y no solo el primer gen estuviesen a su vez silenciados traduccionalmente. En el caso del operón *dszEABC* este silenciamiento se ha demostrado en tres de los 4 genes del operón, incluyendo el último gen (figura 13), lo que demuestra que todos los genes del operón están realmente silenciados. Aunque en principio, la traducción de los distintos cistrones en operones policistrónicos es independiente ya que cada cistrón tiene su propio sitio de unión de ribosomas, es un hecho conocido que la traducción de distintos cistrones muy cercanos o solapantes en un operón pueden estar acoplados traduccionalmente, habiendo demostrado personal de nuestro grupo de investigación que al menos en algunos casos, el cistrón distal ha de ser traducido por la misma subunidad pequeña del ribosoma que ha traducido previamente el cistrón proximal (Govantes *et al.*, 1998). La evidente falta de traducción de estos genes cuando se transcriben de forma “inespecífica” a partir de secuencias aguas arriba, unido a la cercanía o incluso solapamiento entre ellos, indican que estos genes están acoplados traduccionalmente y que el acoplamiento traduccional, combinado con el silenciamiento traduccional del gen proximal, constituye una forma eficaz de evitar la expresión de todos los genes del operón, haciendo que su expresión sea muy precisa, solo desde su propio promotor regulado y constituyendo un mecanismo de regulación que no había sido descrito, hasta el momento, en otros operones de degradación de DBT.

DszR, al igual que SfnR, con la que presenta gran similitud, carece de dominio regulador N-terminal. En la gran mayoría de los casos, el dominio regulador de estos reguladores es de control negativo de su actividad, al demostrarse que mutantes de delección de ese dominio generan activadores transcripcionales con actividad constitutiva. De hecho, existen activadores naturales que carecen de este dominio, los incluidos dentro del grupo IV de activadores transcripcionales, que son constitutivamente activos, aunque su función puede ser controlada por interacción con otra proteína sensora que los inactiva. (Bush & Dixon, 2012). En el caso de SfnR, no se ha descrito ninguna proteína que cumpla esa función, ni tampoco en el caso de DszR se ha podido encontrar en el inserto metagenómico de UPO21 ni UPO112 ninguna proteína que pudiera ejercer ese efecto.

Efectivamente, la actividad de DszR es independiente de la presencia de DBT (figura 18), tal y como se ha descrito para SfnR2 de *P. aeruginosa*, en cuyo caso no es necesaria la presencia de DMS, o sus derivados (Lundgren *et al.*, 2019); sin embargo, al contrario que SfnR, la función activadora de DszR_{UPO21} parecía no estar reprimida por

sulfato. Esto podría deberse a que, en este caso, DszR se estuviera expresando desde el promotor P_{tac} , cuyo nivel basal puede ser suficiente para saturar el sistema, ya que, en condiciones naturales, *dszR* debe expresarse muy poco al tratarse de un gen regulador. De hecho, al testar la expresión de los genes *dsz* en diferentes fuentes de azufre, pero expresados por sí mismos (solo puede hacerse con UPO112), (figura 23), podemos observar que efectivamente la expresión del operón *dsz* está regulada en función de la fuente de azufre, y se asocia a la presencia de fuentes de azufre limitantes, estando reprimida por sulfato, al igual que todos los otros genes *dsz* descritos hasta la fecha (Mohebbi & Ball, 2008) y de forma similar a los genes *sfn* de *Pseudomonas* (Kouzuma, *et al.*, 2008, Lundgren *et al.*, 2019).

Puesto que la transcripción de *dszR*, está a su vez reprimida por sulfato (figura 24, panel derecho) y que la transcripción de *dszE* cuando *dszR* se expresa desde un promotor heterólogo es independiente de la fuente de azufre (figura 30 y 31), la regulación transcripcional de *dszEABC* en función de la fuente de azufre, parece ser ejercida simplemente por el control transcripcional de *dszR* dependiente de la fuente de azufre. La transcripción regulada de *dszR*, a su vez, se activaría por CysB, que es el regulador principal de la respuesta a limitación de azufre tanto en *E. coli* como en *Pseudomonas*, y que podría unirse a la región aguas arriba del operón *dszHR*, en la que se localizan zonas ricas en AT, al igual que sucede en la región promotora de *sfnECR*, (Kouzuma *et al.*, 2008). Según este modelo, que se esquematiza en la figura 37 la regulación de los genes *dsz*, sería, por tanto, muy similar a la propuesta para los de desulfuración de sulfonatos alifáticos en *Pseudomonas*, en la cual CysB activa la transcripción de *sfnR*, y SfnR activa la de los genes *sfn*, en un sistema de regulación en cascada, en el que la fuente de azufre solo regula el primer nivel de activación.

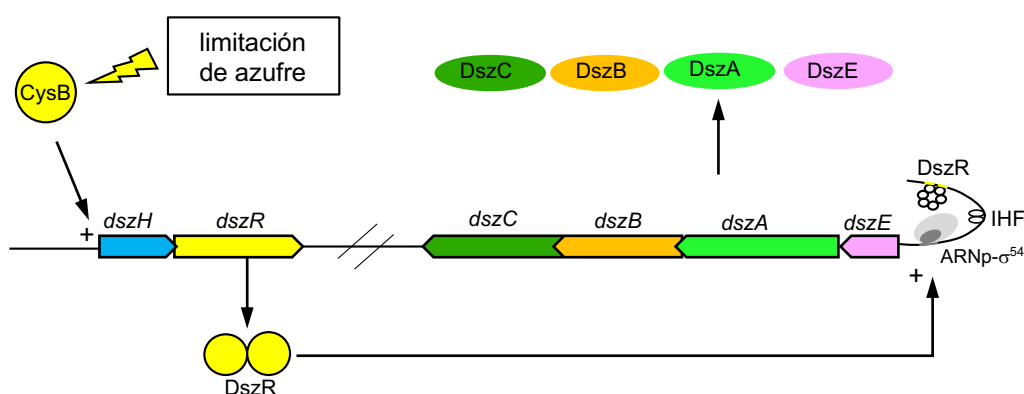


Figura 37: Representación esquemática del modelo propuesto de regulación transcripcional del operón *dszEABC*. En respuesta a limitación de azufre, CysB se une a la región aguas arriba de *dszHR*, activando su expresión. Una vez traducida DszR, ésta se une a sus secuencias de unión

en la región aguas arriba de *dszE*, activando la transcripción del operón *dszEABC* desde el promotor dependiente de σ^{54} , con la ayuda de IHF. La configuración oligomérica de DszR es hipotética.

La similitud estructural y de acción de DszR a los activadores tipo SfnR encontrados en *Pseudomonas* sugieren que DszR podría haber sido reclutado desde otras rutas de degradación de otros compuestos organosulfurados reguladas en respuesta a limitación de azufre, para regular los genes de degradación de DBT, sin necesidad de que la proteína activadora haya tenido que evolucionar para adaptarse a su nuevo operón diana, simplemente, reclutando una región promotora diana de DszR regulable por azufre para el operón *dszEABC*.

IV. Función de DszH.

Teniendo en cuenta que UPO112 se expresa por sí mismo en *E. coli*, pero no en UPO21, y que la transcripción tanto de *dszR* como de *dszEABC* se asocia a fuentes limitantes de azufre (figura 24), es necesario prestar atención al papel de *dszH*, que es el principal elemento diferenciador entre ambos fósmidos. Considerando que en la región entre el -34 y el -112 respecto al ATG de *dszH*, hay un alto contenido en secuencias AT, las cuales están definidas como zonas de unión de CysB en *sfnECR* de *Pseudomonas fluorescens* (Kouzuma *et al.*, 2008), es posible que CysB se una a la región promotora de *dszH* activando a su vez la transcripción de *dszR* como parte del operon *dszHR*. Sin embargo, el gen *dszH* en sí mismo debe ser importante para la expresión de *dszEABC* ya que las dos delecciones internas en fase en *dszH* eliminan por completo la expresión del operón *dszEABC* cuando *dszHR* se transcribe desde su propio promotor, o con niveles basales desde *P_{sal}*, (figura 27), independientemente de que se elimine el promotor de *dszHR* (figura 30).

Dado que la alfa beta hidrolasa codificada por *dszH* no presenta homología con las proteínas bien caracterizadas de las bases de datos, y que su identidad con la proteína codificada por el gen aguas arriba de *dszR* en UPO21 es muy baja, resulta difícil inferir su función. DszH podría tener algún papel en la regulación de la expresión dependiente de fuente de azufre, pudiendo regular los niveles de DszR, o pudiendo ser el responsable de detectar la señal de limitación de azufre.

Pero, teniendo en cuenta que DszR puede activar la transcripción de *dszEABC* independientemente de la fuente de azufre si se transcribe en suficiente cantidad desde

un promotor heterólogo (figura 30 y 31), (incluso con el nivel basal de P_{tac}) haciendo a DszH prescindible, DszH no debe ser el responsable de detectar la fuente de azufre.

Si consideramos que hay muy poca diferencia en los niveles transcripcionales de *dszR* cuando *dszH* es silvestre o si contiene la delección (figura 28), el efecto de DszH sobre la función de DszR, debe ser postranscripcional.

Por otra parte, también hay que tener en cuenta que la función de DszH no se puede complementar en *trans* (figura 29); por tanto, lo que es importante para la función de DszR no es la proteína DszH en sí misma, sino que debe ser en realidad la traducción de su gen completo, descartando así la posibilidad de que DszR necesite la acción de DszH para ser activa, de que pudiera estabilizar a DszR o ser necesaria para su función activadora. Esto concuerda además con el hecho de que la transcripción *in vitro* de *dszE* (figura 22), sea posible sin necesidad de adicionar DszH, o con el hecho de que UPO21 no tenga *dszH*.

En conclusión, y puesto que *dszH* se encuentra solapado con *dszR* 4 pb, por lo que muy probablemente existirá acoplamiento traduccional entre ellos, es de suponer que la traducción de *dszH* sería necesaria para una correcta traducción de *dszR*, circunscribiéndose el papel del gen ayudante a mejorar la traducibilidad de DszR (de ahí el nombre del gen, *H* de *helper*).

Sin embargo, las delecciones generadas en *dszH* son delecciones en fase, de forma que la traducción del segundo gen, *dszR*, debería seguir acoplada en los mutantes de delección, pero, como se ha comentado, la traducción de *dszR* y la transcripción de *dszE* son posibles en el mutante de delección solo cuando *dszR* se transcribe a mayor nivel desde un promotor heterólogo (figura 27, 30 y 31). Existen precedentes en los que la traducción del extremo 3' del cistron proximal no es suficiente para conseguir un acoplamiento traduccional eficaz, identificando regiones mas arriba del cistron proximal que han de ser traducidas para tener una traducción eficaz del gen distal (Little et al, 1989), posiblemente por formación de pseudonudos, en un mecanismo similar al descrito por Asano *et al.*, 1991, o por formación de estructuras secundarias alternativas asociadas al proceso de traducción del primer cistron que puedan facilitar que el RBS del gen segundo cistron pueda mantenerse mas frecuentemente en su forma accesible por el ribosoma (Rex et al, 1994). No se ha realizado un estudio detallado de la implicación de la secuencia de *dszH* en la regulación, por lo que sólo podemos especular con la idea de que algunas de las secuencias que se eliminan con ambas delecciones, puedan tener un papel en la estabilización del inicio de la traducción de *dszR* en una configuración abierta que facilite su traducción.

Por otra parte, es importante tener en cuenta, que se han tratado de conseguir niveles intermedios de expresión de *dszE*, cuando *dszR* se expresa ectópicamente en

ausencia de *dszH* (con diferentes cantidades de inductores y diferentes tiempos de inducción), pero no se ha conseguido. Por tanto, no se puede descartar con absoluta certeza que DszR tenga un papel en la respuesta a fuente de azufre, ya que el sistema de expresión puede estar saturado al inducirlo ectópicamente, y que la pérdida de la regulación por azufre en los plásmidos con la delección P_{dszHR} sea en realidad un artefacto. Puesto que DszR es un regulador, y por tanto necesario en muy poca cantidad, sería posible que, superando un pequeño umbral mínimo, el sistema se desregulase por completo y se hiciera independiente de DszH. Sin embargo, esta posibilidad no es congruente con el hecho de que DszR por sí mismo es activo, como lo son los reguladores tipo IV, que tienen actividad constitutiva, por lo que cualquier control de su actividad ejercido por DszH debería ser inhibitorio de su actividad y, por tanto, las delecciones deberían resultar en una actividad DszR constitutiva.

V. Casetes catabólicos de biodesulfuración de DBT.

Puesto que la construcción de casetes catabólicos para ser expresados en diferentes estirpes es una buena aproximación para conseguir aplicar la BDS industrialmente (Martínez *et al.*, 2016), y dado que los genes *dsz* aislados de la metagenoteca se expresan por sí mismos en un fondo genético diferente al original, y que además de degradar DBT, también pueden degradar 4,6-dimetil-dibenzotiofeno (figura 4), se consideró que podían ser interesantes para un uso industrial. Por tanto, se han construido 2 casetes en vectores de amplio espectro, que pueden ser movilizados por conjugación a distintas bacterias Gram negativas con características de interés para un proceso industrial de biodesulfuración.

Uno de los casetes, contiene los genes *dsz* reguladores y catabólicos por lo que podría expresarse por sí mismo. La incorporación de este casete permite crecer a las 3 estirpes ensayadas utilizando DBT, siendo el crecimiento más eficiente y produciendo más 2-HBF en la estirpe de *Pseudomonas* DOT-T1E. La expresión de los genes *dsz* en las tres estirpes está reprimida por sulfato (figura 35, tabla 11), lo cual concuerda con que los propios genes reguladores del casete estén expresándose en estos fondos genéticos de forma regulada por CysB, que es el regulador principal de la respuesta a limitación de azufre, tanto en *E. coli*, como en *Pseudomonas*.

Un aspecto destacable para el que no tenemos una explicación evidente es el hecho de que la estirpe de *E. coli* MPO554, crezca más lento en DBT, y alcance menos densidad óptica con este vector que con el fósido original (figura 3). De hecho, la cantidad de 2-HBF que produjo indicó que no había agotado el DBT. Una posibilidad es un efecto deletéreo inespecífico del vector en el que se han subclonado los genes *dsz*,

ya que en términos generales estos vectores reducen la velocidad de crecimiento, aunque tampoco se puede descartar la posibilidad de la biodesulfuración sea más eficaz en el contexto genómico original, o que incluso algún gen presente en el fósido pudiera de alguna forma coadyuvar en la biodesulfuración. Para descartar esta opción, habría que subclonar los genes *dsz* en el vector vacío del fósido original y compararlo a UPO112. Sin embargo, esta explicación no parece realmente muy probable, ya que los niveles transcripcionales de *dszE* en UPO112 y en el subclón que contiene *dszHR_{wt}* (pMPO1552), son similares (figuras 23 y 25).

Además, se ha construido un segundo casete, carente de los genes reguladores, en cuyo caso los genes catabólicos se expresan a partir del promotor P_{alkB} , regulado por *alkS*, sistema de expresión que es inducible por alcanos. Aunque este casete, igualmente ha permitido crecer a las estirpes hospedadoras en DBT (figura 36), solo se ha podido observar un efecto positivo del inductor no metabolizable, DCPK en una de las estirpes (MPO406). Puesto que su presencia ha dificultado el crecimiento en todos los casos salvo en éste, incluyendo los cultivos en medio mínimo con sulfato y DBT como fuentes de azufre, en los cuales la adición de DCPK también produjo un crecimiento menor (datos nos mostrados), podemos suponer que este inductor tiene un efecto tóxico.

Dado que el nivel basal de este sistema de expresión, que está descrito como “considerable” (Arce-Rodríguez *et al.*, 2020), ha sido suficiente para expresar los genes *dszEABC*, y permitir el crecimiento en DBT en todas las estirpes, el casete puede ser de gran utilidad, ya que podría emplearse en procesos industriales sin necesidad de adicionar inductor.

Los vectores se han testado en tres estirpes del género *Pseudomonas*, que es uno de los mejores candidatos para ser empleado como hospedador porque posee muchas características que permitirían realizar la BDS de forma más eficiente (Gallardo *et al.*, 1997). En el caso de la estirpe DOT-T1E, la capacidad de tolerar solventes orgánicos es una característica de especial interés, ya que las estirpes tolerantes a solventes orgánicos se pueden emplear en sistemas bifásicos, los cuales tienen mayor eficiencia de BDS que los sistemas acuosos (Derikvand *et al.*, 2015). En el caso de la estirpe MPO406, la capacidad de formar biofilms extensos también es de gran interés, porque el biofilm permite la inmovilización del biocatalizador a un sustrato sólido y proporciona una mayor protección frente a las condiciones de stress que el cultivo planctónico, habiéndose demostrado recientemente (Dorado-Morales *et al.*, 2021) que la formación de biofilm aumenta la capacidad de desulfuración de DBT.

Por otra parte, se ha comprobado que el 2-HBF, que supone un gran problema puesto que causa inhibición sobre las enzimas Dsz (Abin-Fuentes *et al.*, 2013) y además es bactericida, (K. Czechowska *et al.*, 2013), no causa un gran efecto tóxico en las estirpes testadas, lo que confirma que la elección de las cepas hospedadoras es acertada (figura 33), y evita tener que incorporar al sistema los genes de degradación de 2-HBF, *hbpCAD* (Jaspers *et al.*, 2001), posibilidad que se valoró inicialmente, como alternativa, para retirar el 2-HBF del medio y aumentar la eficiencia de la BDS, si el efecto producido por el 2-HBF hubiera sido más dramático.

Puesto que se ha demostrado que los casetes construidos funcionan, ya que permiten biodesulfurar DBT a las estirpes hospedadoras (figura 35 y 36), en un paso posterior estas construcciones se podrían trasladar a un vector transponible, por ejemplo, pBAM1 (Martínez-García *et al.*, 2011), para insertarlo en el genoma de diferentes estirpes, y hacer así el sistema más estable.

CONCLUSIONES

1. UPO112 proporciona a *E. coli* un crecimiento muy eficiente utilizando DBT como única fuente de azufre, llegando a fase estacionaria en menos de 48 horas.
2. UPO112 contiene un operón de degradación *dszEABC* similar al de UPO21, así como el gen regulador del mismo, *dszR*.
3. Los genes *dszE* y *dszR* deben haber sido reclutados a las cercanías de un operón *dszABC* desde sistemas de degradación de sulfonatos alifáticos, fusionándose *dszE* y su promotor con el propio operón *dsz*.
4. Los operones *dsz* de UPO21 y UPO112, deben haberse transmitido por transferencia horizontal, insertándose en genomas muy similares al de *Candidatus* Macondimonas diazotrophica como una isla genómica.
5. El operón *dszEABC* está sujeto a regulación postranscripcional mediante un sistema de aislamiento que, en combinación con el acoplamiento traduccional, impide la traducción de todos los genes del operón cuando se transcriben espúriamente desde promotores aguas arriba de su propio promotor.
6. Los genes *dsz* de UPO112 son capaces de expresarse por sí mismos en *E. coli*, y su expresión está regulada en función de la fuente de azufre, siendo reprimidos por sulfato.
7. La regulación de la transcripción del operón *dszEABC* debe ser dependiente de CysB, el cual, en respuesta a limitación de azufre, activaría la transcripción de *dszR*, y éste a su vez la del operón *dszEABC*.
8. DszR es el primer regulador de genes de biodegradación de DBT caracterizado en bacterias Gram negativas. Se trata de un activador transcripcional de promotores dependientes de sigma⁵⁴, que requiere también el concurso de la proteína IHF para ejercer su función activadora.
9. UPO112, pero no UPO21 contiene un gen denominado *dszH*, localizado aguas arriba de *dszR*, que es esencial para la expresión propia del operón *dszEABC*, y cuya función debe ser mejorar la traducibilidad de *dszR*.
10. DszR consigue activar la transcripción del operón *dszEABC* sin necesidad de *dszH*, si se expresa a un nivel suficiente.
11. Los dos casetes génicos construidos proporcionan la capacidad de biodesulfurar DBT a todas las estirpes ensayadas, que pueden entonces utilizarlo como única fuente de azufre.

ANEXO

Secuencia de aminoácidos de las proteínas Dsz de UPO112.

DszH_{UPO112} 216aa

MQSILAQQINALQAKILIVPGKGNSSESGHWQSLLEARLPDARRVEQANWQQPSLLPWAERIAAAREHDGPVLVVAHS
FGCLATVQALTVLGAPIAATLLVAPADPDRFAIDPTWLQSAALPGVHRMVISNNDPWLSPARAHALARAWKIPALSLGN
AGHINVAAGFGAWPRGDVLLAKLLARLIDRSAAGVSLLPADSPSANTPFPSQFARQGIALL

DszR_{UPO112} 377aa

MKFKTDPPAEPTLSLDIPHGTPATRTAPAVLTNLKLAQQPHSIRASALVFADPRSQELLDYLERISPSDANVLIQGE
TGTGKELVARHVHELSTGLQKGPFFVAINCAALNENLLDAELFGYRKGAFTGADSDRAGWFETAQGGTLFLDEIGDLPLS
AQVKLLRVLQQREVVRVGARTPVSVNFRLIAATNVNLKTAVEAGHFRPDLYRSLVAEVRPLPLRERPGDIAPLVSHF
LGIYQRLGYRHANIQPQALDQLQHSWPGNIRELENNVHHTLLVCPDTEIRPEDIHLPERPQRPGESVTTPEARLQQ
VFRELMASHSTADLHAVSERILLETAYQHSGDNQIQAAKALGISRNVLRERLQRLGIIPYPPSRRS

DszC_{UPO112} 412aa

MSQSAENLYVHQNDPVEVARQLAAKWQATAVDRDRAGGSATAERQDLRDSGLLSLLIPKEYGGWADWPTTIEVIREI
AKVDGSLGHLFGYHTGSAPMIELFGSTEQKERLYQQLAQNNWWTGNASSENNSHVLEWKVSATPTEDGGYLLNGTKHF
CSGAKGSDLLFVFGVIEESPOKGAIVSVAIPTGREGVHINDDNAIGMRQTDSGSTEFHNVKVYPDEVLGAPNSFIL
AFMGSKRGSWTPMVQLIFSNVYLGIAHGALEAAAREYTRTQTRPWTAGVKKATEDPYVIRTYGEFAIELQGADAAAR
EAARLLQQVWEKGDALTPKDRGELMVKVSQVKGALSTKVALDITSRIFEVIGSRGTHPKYGFDRFWRNVRTSLHDPVA
YKIVEVGNHALNGRYPVPGFTS

DszB_{UPO112} 371aa

MASRPAPQSDRLNSTQAPAAATLTYSNCPVPNALLTALSSGILSEAGIQDLILSGKQGAHVHTYDQPAYTRYGGEIPPL
LSEGLRAPGRTRLGITPVLGRKGYFVRGDSPIQKPADLAGRRVGVSSASATRILRGELGDYLELDPWRQTLVALGTWE
ARALLHTEAGGLSINDVELVRIESPGVDPPERLQAASSLKGADLFPDVAGQQAILEQGAEALFAWLPWAAELEA
RIGARPVIDLGQDVRNAYASVWTVSSELAARQPEVVQRLVDAAVDAGLWAQENGDAVTRLHAENLGVTENGIRQGFGP
DFHQHLVPRLDTEAFAILARTQQFLLDKLLHEPVTLDLWAAPEFLNNSLKRNLTRKIA

DszA_{UPO112} 464aa

MSEPKQLHLAGFFSAGNVTHAHGAWRHARATNGFLTKEYYQHIARTLERGKFDLLFLPDGLAVEDGYGDNLETGVRYG
GQGAVALPEAAVIATMAAVTERLGLGATISTTYPPYHVARVFASLDQLSGGRISWNVVTSLNNAEARNFGVDQHLEH
DSRYDRADEFLEAVRKLWNSWDADALVLDKANGQFADPSKVRYVDHRGEWLNVRGPLQVPRSPQGEPIVLQAGLSDRG
KRFAGRWAEGVFSVAPNLAQMOSNYRDIKAQVKAAGRDPDQTKIFTAVMPVLGETEAIARERLEYLNSLVHPEVGLST
LSSHTGINLAVYPLDTLMTFVLRDLKERNVPTPLHMFMTAMQANALTLGEMGRRYGTNVAFVPQWVGTGEQIADQIET
HFKAGAADGFIISPAFLPGAYDEFVDQVVPILQARGLFRTEYEGQTLRDHLGLREPQPVGENTWQAAQRHKAIA

DszE_{UPO112} 192 aa

MPHILILNGSPSAESRTGRLLDHFAEQSLDGWSVNQITLRTLPAQALLSADFALPLVSARLAVEEADAIVVATPVY
KGAYTGLLKAFIDTLPHRALVDKSVLPPIASGGTLAHYGVIDHALKPILAALGATYTLPGIFVLDSDLKTPQGVHGFNA
EIVDRLANARLALRRLHRLAADAVPEADTHLRQYGS

BIBLIOGRAFÍA

Abbasian F, Lockington R, Megharaj M, Naidu R. Identification of a new operon involved in desulfurization of dibenzothiophenes using a metagenomic study and cloning and functional analysis of the genes. *Enzyme Microb Technol.* 2016 Jun;87-88:24-8. doi: 10.1016/j.enzmictec.2016.02.009.

Abin-Fuentes A, Mohamed Mel-S, Wang DI, Prather KL. Exploring the mechanism of biocatalyst inhibition in microbial desulfurization. *Appl Environ Microbiol.* 2013 Dec;79(24):7807-17. doi: 10.1128/AEM.02696-13.

Abril MA, Michan C, Timmis KN, Ramos JL. Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. *J Bacteriol.* 1989 Dec;171(12):6782-90. doi: 10.1128/jb.171.12.6782-6790.1989. Erratum in: *J Bacteriol* 1990 Jun;172(6):3534.

Amador CI, Canosa I, Govantes F, Santero E. Lack of CbrB in *Pseudomonas putida* affects not only amino acids metabolism but also different stress responses and biofilm development. *Environ Microbiol.* 2010 Jun;12(6):1748-61. doi: 10.1111/j.1462-2920.2010.02254.x.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W & Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* (1997), 25(17):3389-402. doi: 10.1093/nar/25.17.3389.

Alves L, Melo M, Mendonça D, Simões F, Matos J, Tenreiro R, F.M. Gírio FM Sequencing, cloning and expression of the dsz genes required for dibenzothiophene sulfone desulfurization from *Gordonia alkanivorans* strain 1B. *Enzyme and Microbial Technology*, Volume 40, Issue 6, 2007, Pages 1598-1603, ISSN 0141-0229. doi.org/10.1016/j.enzmictec.2006.11.008

Arce-Rodríguez A, Benedetti I, Silva-Rocha R, de Lorenzo V. Standardization of inducer-activated broad host range expression modules: Debugging and refactoring an alkane-responsive *AlkS/P_{alkB}* device bioRxiv 2020.12.26.424440; doi.org/10.1101/2020.12.26.424440 (Preprint).

Asano K, Kato A, Moriwaki H, Hama C, Shiba K, Mizobuchi K. Positive and negative regulations of plasmid ColIb-P9 repZ gene expression at the translational level. *J Biol Chem.* 1991 Feb 25;266(6):3774-81.

Bagdasarian M, Lurz R, Rückert B, Franklin FC, Bagdasarian MM, Frey J, Timmis KN. Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene.* 1981 Dec;16(1-3):237-47. doi: 10.1016/0378-1119(81)90080-9.

Beloqui A, Polaina J, Vieites JM, Reyes-Duarte D, Torres R, Golyshina OV, Chernikova TN, Waliczek A, Aharoni A, Yakimov MM, Timmis KN, Golyshin PN, Ferrer M. Novel hybrid esterase-haloacid dehalogenase enzyme. *Chembiochem.* 2010 Sep 24;11(14):1975-8. doi: 10.1002/cbic.201000258.

Bordoloi, N. K., Rai, S. K., Chaudhuri, M. K., & Mukherjee, A. K. (2014). Deep-desulfurization of dibenzothiophene and its derivatives present in diesel oil by a newly isolated bacterium *Achromobacter* sp. to reduce the environmental pollution from fossil fuel combustion. *Fuel Processing Technology*, 119, 236–244. doi:10.1016/j.fuproc.2013.10.014

Bordoloi NK, Bhagowati P, Chaudhuri MK, Mukherjee AK. Proteomics and Metabolomics Analyses to Elucidate the Desulfurization Pathway of *Chelatococcus* sp. *PLoS One.* 2016 Apr 21;11(4):e0153547. doi: 10.1371/journal.pone.0153547.

Brennerova MV, Josefiova J, Brenner V, Pieper DH, Junca H. Metagenomics reveals diversity and abundance of meta-cleavage pathways in microbial communities from soil highly contaminated with jet fuel under air-sparging bioremediation. *Environ Microbiol.* 2009 Sep;11(9):2216-27. doi: 10.1111/j.1462-2920.2009.01943.x.

Bush M, Dixon R. The role of bacterial enhancer binding proteins as specialized activators of σ^{54} -dependent transcription. *Microbiol Mol Biol Rev.* 2012 Sep;76(3):497-529. doi: 10.1128/MMBR.00006-12.

Bykowski T, van der Ploeg JR, Iwanicka-Nowicka R, Hryniewicz MM. The switch from inorganic to organic sulphur assimilation in *Escherichia coli*: adenosine 5'-phosphosulphate (APS) as a signalling molecule for sulphate excess. *Mol Microbiol.* 2002 Mar;43(5):1347-58. doi: 10.1046/j.1365-2958.2002.02846.x.

Cases I, de Lorenzo V. The black cat/white cat principle of signal integration in bacterial promoters. *EMBO J.* 2001 Jan 15;20(1-2):1-11. doi: 10.1093/emboj/20.1.1.

Chung CT, Niemela SL & Miller RH. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci U S A* (1989), 86(7):2172-5. doi: 10.1073/pnas.86.7.2172.

Czechowska K, Reimmann C, van der Meer JR. Characterization of a MexAB-OprM efflux system necessary for productive metabolism of *Pseudomonas azelaica* HBP1 on 2-hydroxybiphenyl. *Front Microbiol.* 2013 Jul 19;4:203. doi: 10.3389/fmicb.2013.00203.

Davison J, Brunel F, Phanopoulos A, Prozzi D, Terpstra P. Cloning and sequencing of *Pseudomonas* genes determining sodium dodecyl sulfate biodegradation. *Gene.* 1992 May 1;114(1):19-24. doi: 10.1016/0378-1119(92)90702-q.

Derikvand P, Etemadifar Z, Saber H. Sulfur Removal from Dibenzothiophene by Newly Isolated *Paenibacillus validus* Strain PD2 and Process Optimization in Aqueous and Biphasic (Model-Oil) Systems. *Pol J Microbiol.* 2015;64(1):47-54. PMID: 26094315.

Di Gregorio S, Zocca C, Sidler S, Toffanin A, Lizzari D, Vallini G. Identification of two new sets of genes for dibenzothiophene transformation in *Burkholderia* sp. DBT1. *Biodegradation.* 2004 Apr;15(2):111-23. doi: 10.1023/b:biod.0000015624.52954.b6.

Dorado-Morales P, Martínez I, Rivero-Buceta V, Díaz E, Bähre H, Lasa I, Solano C. Elevated c-di-GMP levels promote biofilm formation and biodesulfurization capacity of *Rhodococcus erythropolis*. *Microb Biotechnol.* 2021 May;14(3):923-937. doi: 10.1111/1751-7915.13689.

Elliott T, Geiduschek EP. Defining a bacteriophage T4 late promoter: absence of a "-35" region. *Cell.* 1984 Jan;36(1):211-9. doi: 10.1016/0092-8674(84)90091-6. P

Endoh T, Habe H, Yoshida T, Nojiri H, Omori T. A CysB-regulated and sigma54-dependent regulator, SfnR, is essential for dimethyl sulfone metabolism of *Pseudomonas putida* strain DS1. *Microbiology (Reading).* 2003a Apr;149(Pt 4):991-1000. doi: 10.1099/mic.0.26031-0.

Endoh T, Kasuga K, Horinouchi M, Yoshida T, Habe H, Nojiri H, Omori T. Characterization and identification of genes essential for dimethyl sulfide utilization in *Pseudomonas putida* strain DS1. *Appl Microbiol Biotechnol.* 2003b Jul;62(1):83-91. doi: 10.1007/s00253-003-1233-7.

Endoh T, Habe H, Nojiri H, Yamane H, Omori T. The sigma54-dependent transcriptional activator SfnR regulates the expression of the *Pseudomonas putida* sfnFG operon responsible for dimethyl sulphone utilization. *Mol Microbiol.* 2005 Feb;55(3):897-911. doi: 10.1111/j.1365-2958.2004.04431.x.

Feng, S., H. Yang, X. Zhan, and W. Wang. 2016. Enhancement of dibenzothiophene biodesulfurization by weakening the feedback inhibition effects based on a systematic understanding of the biodesulfurization mechanism by *Gordonia* sp. through the potential "4S" pathway. *RSC Adv.* 6:82872–81. 10.1039/C6RA14459D

Fernández-Arrojo L, Guazzaroni ME, López-Cortés N, Beloqui A, Ferrer M. Metagenomic era for biocatalyst identification. *Curr Opin Biotechnol*. 2010 Dec;21(6):725-33. doi: 10.1016/j.copbio.2010.09.006.

Ferrer M, Beloqui A, Timmis KN, Golyshin PN. Metagenomics for mining new genetic resources of microbial communities. *J Mol Microbiol Biotechnol*. 2009;16(1-2):109-23. doi: 10.1159/000142898.

Figuroa-Gonzalez PA, Bornemann TLV, Adam PS, Plewka J, Révész F, von Hagen CA, Táncsics A, Probst AJ. Saccharibacteria as Organic Carbon Sinks in Hydrocarbon-Fueled Communities. *Front Microbiol*. 2020 Dec 23;11:587782. doi: 10.3389/fmicb.2020.587782.

Figurski DH, Helinski DR. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci U S A*. 1979 Apr;76(4):1648-52. doi: 10.1073/pnas.76.4.1648.

Filiseti L, Fontecave M, Niviere V. Mechanism and substrate specificity of the flavin reductase ActVB from *Streptomyces coelicolor*. *J Biol Chem*. 2003 Jan 3;278(1):296-303. doi: 10.1074/jbc.M209689200.

Furuya T, Takahashi S, Iwasaki Y, Ishii Y, Kino K, Kirimura K. Gene cloning and characterization of *Mycobacterium phlei* flavin reductase involved in dibenzothiophene desulfurization. *J Biosci Bioeng*. 2005 Jun;99(6):577-85. doi: 10.1263/jbb.99.577.

Galán B, Díaz E, García JL. Enhancing desulphurization by engineering a flavin reductase-encoding gene cassette in recombinant biocatalysts. *Environ Microbiol*. 2000 Dec;2(6):687-94. doi: 10.1046/j.1462-2920.2000.00151.x.

Gallardo ME, Ferrández A, De Lorenzo V, García JL, Díaz E. Designing recombinant *Pseudomonas* strains to enhance biodesulfurization. *J Bacteriol*. 1997 Nov;179(22):7156-60. doi: 10.1128/jb.179.22.7156-7160.1997.

Gao F, Danson AE, Ye F, Jovanovic M, Buck M, Zhang X. Bacterial Enhancer Binding Proteins-AAA+ Proteins in Transcription Activation. *Biomolecules*. 2020 Feb 25;10(3):351. doi: 10.3390/biom10030351.

García-González V, Govantes F, Porrúa O, Santero E. Regulation of the *Pseudomonas* sp. strain ADP cyanuric acid degradation operon. *J Bacteriol*. 2005 Jan;187(1):155-67. doi: 10.1128/JB.187.1.155-167.2005.

García-Mauriño SM, Pérez-Martínez I, Amador CI, Canosa I, Santero E. Transcriptional activation of the *CrcZ* and *CrcY* regulatory RNAs by the *CbrB* response regulator in *Pseudomonas putida*. *Mol Microbiol*. 2013 Jul;89(1):189-205. doi: 10.1111/mmi.12270.

Geronimo I, Nigam SR, Payne CM. Desulfination by 2'-hydroxybiphenyl-2-sulfinate desulfinate proceeds via electrophilic aromatic substitution by the cysteine-27 proton. *Chem Sci*. 2017 Jul 1;8(7):5078-5086. doi: 10.1039/c7sc00496f.

Govantes F, Santero E. Transcription termination within the regulatory *nifLA* operon of *Klebsiella pneumoniae*. *Mol Gen Genet*. 1996 Mar 7;250(4):447-54.

Govantes F, Andújar E, Santero E. Mechanism of translational coupling in the *nifLA* operon of *Klebsiella pneumoniae*. *EMBO J*. 1998 Apr 15;17(8):2368-77. doi: 10.1093/emboj/17.8.2368.

Govantes F, Albrecht JA, Gunsalus RP. Oxygen regulation of the *Escherichia coli* cytochrome d oxidase (*cydAB*) operon: roles of multiple promoters and the *Fnr-1* and *Fnr-2* binding sites. *Mol Microbiol*. 2000 Sep;37(6):1456-69. doi: 10.1046/j.1365-2958.2000.02100.x.

Gray KA, Pogrebinsky OS, Mrachko GT, Xi L, Monticello DJ, Squires CH. Molecular mechanisms of biocatalytic desulfurization of fossil fuels. *Nat Biotechnol.* 1996 Dec;14(13):1705-9. doi: 10.1038/nbt1296-1705.

Gray KA, Mrachko GT, Squires CH. Biodesulfurization of fossil fuels. *Curr Opin Microbiol.* 2003 Jun;6(3):229-35. doi: 10.1016/s1369-5274(03)00065-1.

Habe H, Kouzuma A, Endoh T, Omori T, Yamane H, Nojiri H. Transcriptional regulation of the sulfate-starvation-induced gene *sfnA* by a sigma54-dependent activator of *Pseudomonas putida*. *Microbiology (Reading).* 2007 Sep;153(Pt 9):3091-3098. doi: 10.1099/mic.0.2007/008151-0.

Hanahan D. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol.* 1983 Jun 5;166(4):557-80. doi: 10.1016/s0022-2836(83)80284-8.

Inoue H, Nojima H, Okayama H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene.* 1990 Nov 30;96(1):23-8. doi: 10.1016/0378-1119(90)90336-p

Ishii Y, Konishi J, Okada H, Hirasawa K, Onaka T, Suzuki M. Operon structure and functional analysis of the genes encoding thermophilic desulfurizing enzymes of *Paenibacillus* sp. A11-2. *Biochem Biophys Res Commun.* 2000 Apr 2;270(1):81-8. doi: 10.1006/bbrc.2000.2370.

Ishii Y, Konishi J, Suzuki M, Maruhashi K. Cloning and expression of the gene encoding the thermophilic NAD(P)H-FMN oxidoreductase coupling with the desulfurization enzymes from *Paenibacillus* sp. A11-2. *J Biosci Bioeng.* 2000;90(6):591-9. doi: 10.1263/jbb.90.591.

Ismail W, El Noyal AM, Ramadan AR, Abotalib N. Sulfur source-mediated transcriptional regulation of the *rhlABC* genes involved in biosurfactants production by *Pseudomonas* sp. strain AK6U. *Front Microbiol.* 2014 Aug 14;5:423. doi: 10.3389/fmicb.2014.00423.

Jaspers MC, Schmid A, Sturme MH, Goslings DA, Kohler HP, Roelof Van Der Meer J. Transcriptional organization and dynamic expression of the *hbpCAD* genes, which encode the first three enzymes for 2-hydroxybiphenyl degradation in *Pseudomonas azelaica* HBP1. *J Bacteriol.* 2001 Jan;183(1):270-9. doi: 10.1128/JB.183-1.270-279.2001.

Kahnert A, Vermeij P, Wietek C, James P, Leisinger T, Kertesz MA. The *ssu* locus plays a key role in organosulfur metabolism in *Pseudomonas putida* S-313. *J Bacteriol.* 2000 May;182(10):2869-78. doi: 10.1128/JB.182.10.2869-2878.2000.

Kahnert A, Mirleau P, Wait R, Kertesz MA. The LysR-type regulator SftR is involved in soil survival and sulphate ester metabolism in *Pseudomonas putida*. *Environ Microbiol.* 2002 Apr;4(4):225-37. doi: 10.1046/j.1462-2920.2002.00289.x.

Karthikeyan S, Rodriguez-R LM, Heritier-Robbins P, Kim M, Overholt WA, Gaby JC, Hatt JK, Spain JC, Rosselló-Móra R, Huettel M, Kostka JE, Konstantinidis KT. "*Candidatus Macondimonas diazotrophica*", a novel gammaproteobacterial genus dominating crude-oil-contaminated coastal sediments. *ISME J.* 2019 Aug;13(8):2129-2134. doi: 10.1038/s41396-019-0400-5.

Kawaguchi H, Kobayashi H, Sato K. Metabolic engineering of hydrophobic *Rhodococcus opacus* for biodesulfurization in oil-water biphasic reaction mixtures. *J Biosci Bioeng.* 2012 Mar;113(3):360-6. doi: 10.1016/j.jbiosc.2011.10.017.

Kertesz MA. Riding the sulfur cycle-metabolism of sulfonates and sulfate esters in gram-negative bacteria. *FEMS Microbiol Rev.* 2000 Apr;24(2):135-75. doi: 10.1016/S0168-6445(99)00033-9.

Kilbane, J. J. (1989). Desulfurization of coal: the microbial solution. *Trends in Biotechnology*, 7(4), 97–101. doi:10.1016/0167-7799(89)90007-3

Kilbane JJ 2nd. Microbial biocatalyst developments to upgrade fossil fuels. *Curr Opin Biotechnol.* 2006 Jun;17(3):305-14. doi: 10.1016/j.copbio.2006.04.005.

Kilbane, John J. (2016). Biodesulfurization: How to Make it Work?. *Arabian Journal for Science and Engineering*, 42(1), 1–9. doi:10.1007/s13369-016-2269-1

Koch DJ, Rückert C, Rey DA, Mix A, Pühler A, Kalinowski J. Role of the *ssu* and *seu* genes of *Corynebacterium glutamicum* ATCC 13032 in utilization of sulfonates and sulfonate esters as sulfur sources. *Appl Environ Microbiol.* 2005 Oct;71(10):6104-14. doi: 10.1128/AEM.71.10.6104-6114.2005.

Kouzuma A, Endoh T, Omori T, Nojiri H, Yamane H, Habe H. Transcription factors CysB and SfnR constitute the hierarchical regulatory system for the sulfate starvation response in *Pseudomonas putida*. *J Bacteriol.* 2008 Jul;190(13):4521-31. doi: 10.1128/JB.00217-08.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol.* 2018 Jun 1;35(6):1547-1549. doi: 10.1093/molbev/msy096.

Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 1970 Aug 15;227(5259):680-5. doi: 10.1038/227680a0.

Levy L, Anavy L, Solomon O, Cohen R, Brunwasser-Meirom M, Ohayon S, Atar O, Goldberg S, Yakhini Z, Amit R. A Synthetic Oligo Library and Sequencing Approach Reveals an Insulation Mechanism Encoded within Bacterial σ^{54} Promoters. *Cell Rep.* 2017 Oct 17;21(3):845-858. doi: 10.1016/j.celrep.2017.09.063.

Leong SA, Ditta GS, Helinski DR. Heme biosynthesis in *Rhizobium*. Identification of a cloned gene coding for delta-aminolevulinic acid synthetase from *Rhizobium meliloti*. *J Biol Chem.* 1982 Aug 10;257(15):8724-30.

Lorenz R, Bernhart SH, Höner Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. ViennaRNA Package 2.0. *Algorithms Mol Biol.* 2011 Nov 24;6:26. doi: 10.1186/1748-7188-6-26.

Lundgren BR, Sarwar Z, Feldman KS, Shoytush JM, Nomura CT. SfnR2 Regulates Dimethyl Sulfide-Related Utilization in *Pseudomonas aeruginosa* PAO1. *J Bacteriol.* 2019 Jan 28;201(4):e00606-18. doi: 10.1128/JB.00606-18.

Li MZ, Squires CH, Monticello DJ, Childs JD. Genetic analysis of the *dsz* promoter and associated regulatory regions of *Rhodococcus erythropolis* IGTS8. *J Bacteriol.* 1996 Nov;178(22):6409-18. doi: 10.1128/jb.178.22.6409-6418.1996.

Li F, Xu P, Feng J, Meng L, Zheng Y, Luo L, Ma C. Microbial desulfurization of gasoline in a *Mycobacterium goodii* X7B immobilized-cell system. *Appl Environ Microbiol.* 2005 Jan;71(1):276-81. doi: 10.1128/AEM.71.1.276-281.2005.

Li GQ, Ma T, Li SS, Li H, Liang FL, Liu RL. Improvement of dibenzothiophene desulfurization activity by removing the gene overlap in the *dsz* operon. *Biosci Biotechnol Biochem.* 2007 Apr;71(4):849-54. doi: 10.1271/bbb.60189.

Li GQ, Li SS, Zhang ML, Wang J, Zhu L, Liang FL, Liu RL, Ma T. Genetic rearrangement strategy for optimizing the dibenzothiophene biodesulfurization pathway in *Rhodococcus erythropolis*. *Appl Environ Microbiol.* 2008 Feb;74(4):971-6. doi: 10.1128/AEM.02319-07.

Li L, Liao Y, Luo Y, Zhang G, Liao X, Zhang W, Zheng S, Han S, Lin Y, Liang S. Improved Efficiency of the Desulfurization of Oil Sulfur Compounds in *Escherichia coli* Using a Combination of Desensitization Engineering and *DszC* Overexpression. *ACS Synth Biol.* 2019 Jun 21;8(6):1441-1451. doi: 10.1021/acssynbio.9b00126.

Little S, Hyde S, Campbell CJ, Lilley RJ, Robinson MK. Translational coupling in the threonine operon of *Escherichia coli* K-12. *J Bacteriol.* 1989 Jun;171(6):3518-22. doi: 10.1128/jb.171.6.3518-3522.1989.

Mandelbaum RT, Wackett LP, Allan DL. Mineralization of the s-triazine ring of atrazine by stable bacterial mixed cultures. *Appl Environ Microbiol.* 1993 Jun;59(6):1695-701. doi: 10.1128/aem.59.6.1695-1701.1993.

Martín-Cabello G, Terrón-González L, Ferrer M, Santero E. Identification of a complete dibenzothiophene biodesulfurization operon and its regulator by functional metagenomics. *Environ Microbiol.* 2020 Jan;22(1):91-106. doi: 10.1111/1462-2920.14823.

Martínez I, Mohamed ME, Rozas D, García JL, Díaz E. Engineering synthetic bacterial consortia for enhanced desulfurization and revalorization of oil sulfur compounds. *Metab Eng.* 2016 May;35:46-54. doi: 10.1016/j.ymben.2016.01.005.

Martínez I, El-Said Mohamed M, Santos VE, García JL, García-Ochoa F, Díaz E. Metabolic and process engineering for biodesulfurization in Gram-negative bacteria. *J Biotechnol.* 2017 Nov 20;262:47-55. doi: 10.1016/j.jbiotec.2017.09.004.

Martínez-García E, Calles B, Arévalo-Rodríguez M, de Lorenzo V. pBAM1: an all-synthetic genetic tool for analysis and construction of complex bacterial phenotypes. *BMC Microbiol.* 2011 Feb 22;11:38. doi: 10.1186/1471-2180-11-38.

Martínez-Pérez O, Moreno-Ruiz E, Floriano B, Santero E. Regulation of tetralin biodegradation and identification of genes essential for expression of the *thn* operons. *J Bacteriol.* 2004 Sep;186(18):6101-9. doi: 10.1128/JB.186.18.6101-6109.2004.

Miller JH. *Experiments in Molecular Genetics*, Cold Spring Harbor: Cold Spring Harbor Laboratory Press (1972).

Mohebbi G, Ball AS. Biocatalytic desulfurization (BDS) of petrodiesel fuels. *Microbiology (Reading)*. 2008 Aug;154(Pt 8):2169-2183. doi: 10.1099/mic.0.2008/017608-0.

Mohebbi, G., Ball, A.S., 2016. Biodesulfurization of diesel fuels – Past, present and future perspectives. *Int. Biodeter. Biodegrad.* 110, 163-180. DOI:10.1016/j.ibiod.2016.03.011

Murarka P, Bagga T, Singh P, Rangra S, Srivastava P. Isolation and identification of a TetR family protein that regulates the biodesulfurization operon. *AMB Express.* 2019a May 24;9(1):71. doi: 10.1186/s13568-019-0801-x.

Murarka P, Srivastava P. Characterization of DNA binding and ligand binding properties of the TetR family protein involved in regulation of *dsz* operon in *Gordonia* sp. IITR100. *Int J Biol Macromol.* 2019b Dec 1;141:671-679. doi: 10.1016/j.ijbiomac.2019.09.009.

Murarka P, Keshav A, Meena BK, Srivastava P. Functional characterization of the transcription regulator WhiB1 from *Gordonia* sp. IITR100. *Microbiology (Reading)*. 2020 Dec;166(12):1181-1190. doi: 10.1099/mic.0.000985.

Nassar HN, Abu Amr SS, El-Gendy NS. Biodesulfurization of refractory sulfur compounds in petro-diesel by a novel hydrocarbon tolerable strain *Paenibacillus glucanolyticus* HN4. *Environ Sci Pollut Res Int.* 2021 Feb;28(7):8102-8116. doi: 10.1007/s11356-020-11090-7.

Noda K, Watanabe K, Maruhashi K. Isolation of the *Pseudomonas aeruginosa* gene affecting uptake of dibenzothiophene in n-tetradecane. *J Biosci Bioeng.* 2003;95(5):504-11.

Noda K, Watanabe K, Maruhashi K. Recombinant *Pseudomonas putida* carrying both the *dsz* and *hcu* genes can desulfurize dibenzothiophene in n-tetradecane. *Biotechnol Lett.* 2003b Jul;25(14):1147-50. doi: 10.1023/a:1024521313249.

Pan, J., Wu, F., Wang, J., Yu, L., Khayyat, N. H., Stark, B. C., and Kilbane, J. J. (2013) Enhancement of desulfurization activity by enzymes of the *Rhodococcus dsz* operon through coexpression of a high sulfur peptide and directed evolution. *Fuel* 112, 385–390. 2013, Pages 385-390, ISSN 0016-2361, doi.org/10.1016/j.fuel.2013.04.065.

Park HS, Kayser KJ, Kwak JH, Kilbane JJ 2nd. Heterologous gene expression in *Thermus thermophilus*: beta-galactosidase, dibenzothiophene monooxygenase, PNB carboxy esterase, 2-aminobiphenyl-2,3-diol dioxygenase, and chloramphenicol acetyl transferase. *J Ind Microbiol Biotechnol*. 2004 May;31(4):189-97. doi: 10.1007/s10295-004-0130-0.

Parveen S, Akhtar N, Ghauri MA, Akhtar K. Conventional genetic manipulation of desulfurizing bacteria and prospects of using CRISPR-Cas systems for enhanced desulfurization activity. *Crit Rev Microbiol*. 2020 May;46(3):300-320. doi: 10.1080/1040841X.2020.1772195.

Peng C, Huang D, Shi Y, Zhang B, Sun L, Li M, Deng X, Wang W. Comparative transcriptomic analysis revealed the key pathways responsible for organic sulfur removal by thermophilic bacterium *Geobacillus thermoglucosidasius* W-2. *Sci Total Environ*. 2019 Aug 1;676:639-650. doi: 10.1016/j.scitotenv.2019.04.328.

Piddington CS, Kovacevich BR, Rambosek J. Sequence and molecular characterization of a DNA region encoding the dibenzothiophene desulfurization operon of *Rhodococcus* sp. strain IGTS8. *Appl Environ Microbiol*. 1995 Feb;61(2):468-75. doi: 10.1128/aem.61.2.468-475.1995.

Platero AI, García-Jaramillo M, Santero E, Govantes F. Transcriptional organization and regulatory elements of a *Pseudomonas* sp. strain ADP operon encoding a LysR-type regulator and a putative solute transport system. *J Bacteriol*. 2012 Dec;194(23):6560-73. doi: 10.1128/JB.01348-12.

Platero AI, López-Sánchez A, Tomás-Gallardo L, Santero E, Govantes F. Mechanism of Antiactivation at the *Pseudomonas* sp. Strain ADP σ N-Dependent PatzT Promoter. *Appl Environ Microbiol*. 2016 Jun 30;82(14):4350-4362. doi: 10.1128/AEM.00906-16.

Porrúa O, García-Jaramillo M, Santero E, Govantes F. The LysR-type regulator AtzR binding site: DNA sequences involved in activation, repression and cyanuric acid-dependent repositioning. *Mol Microbiol*. 2007 Oct;66(2):410-27. doi: 10.1111/j.1365-2958.2007.05927.x.

Porrúa O, García-González V, Santero E, Shingler V, Govantes F. Activation and repression of a sigmaN-dependent promoter naturally lacking upstream activation sequences. *Mol Microbiol*. 2009 Aug;73(3):419-33. doi: 10.1111/j.1365-2958.2009.06779.x.

Ramos JL, Duque E, Huertas MJ, Haïdour A. Isolation and expansion of the catabolic potential of a *Pseudomonas putida* strain able to grow in the presence of high concentrations of aromatic hydrocarbons. *J Bacteriol*. 1995 Jul;177(14):3911-6. doi: 10.1128/jb.177.14.3911-3916.1995.

Rex G, Surin B, Besse G, Schneppe B, McCarthy JE. The mechanism of translational coupling in *Escherichia coli*. Higher order structure in the atpHA mRNA acts as a conformational switch regulating the access of de novo initiating ribosomes. *J Biol Chem*. 1994 Jul 8;269(27):18118-27.

Rojas A, Duque E, Schmid A, Hurtado A, Ramos JL, Segura A. Biotransformation in double-phase systems: physiological responses of *Pseudomonas putida* DOT-T1E to a double phase made of aliphatic alcohols and biosynthesis of substituted catechols. *Appl Environ Microbiol*. 2004 Jun;70(6):3637-43. doi: 10.1128/AEM.70.6.3637-3643.2004.

Rivas-Marín E, Floriano B, Santero E. Genetic dissection of independent and cooperative transcriptional activation by the LysR-type activator ThnR at close divergent promoters. *Sci Rep*. 2016 Apr 18;6:24538. doi: 10.1038/srep24538.

Samal, S.K.. "Leader Sequence." Brenner's Encyclopedia of Genetics (2013): 203–205. doi:10.1016/B978-0-12-374984-0.00850-0

Sambrook, J. y Russell, D. W. (2001) Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Salis HM, Mirsky EA, Voigt CA. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol.* 2009 Oct;27(10):946-50. doi: 10.1038/nbt.1568.

Shingler V. Signal sensory systems that impact σ^{54} -dependent transcription. *FEMS Microbiol Rev.* 2011 May;35(3):425-40. doi: 10.1111/j.1574-6976.2010.00255.x.

Silva-Rocha R, Martínez-García E, Calles B, Chavarría M, Arce-Rodríguez A, de Las Heras A, Páez-Espino AD, Durante-Rodríguez G, Kim J, Nikel PI, Platero R, de Lorenzo V. The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* 2013 Jan;41(Database issue):D666-75. doi: 10.1093/nar/gks1119. Epub 2012 Nov 23. PMID: 23180763; PMCID: PMC3531073.

Stephen D, Jones C, Schofield JP. A rapid method for isolating high quality plasmid DNA suitable for DNA sequencing. *Nucleic Acids Res.* 1990 Dec 25;18(24):7463-4. doi: 10.1093/nar/18.24.7463.

Takada M, Nomura N, Okada H, Nakajima-Kambe T, Nakahara T, Uchiyama H. De-repression and comparison of oil-water separation activity of the dibenzothiophene desulfurizing bacterium, *Mycobacterium* sp. G3. *Biotechnol Lett.* 2005 Jun;27(12):871-4. doi: 10.1007/s10529-005-6721-7.

Tanaka Y, Yoshikawa O, Maruhashi K, Kurane R. The cbs mutant strain of *Rhodococcus erythropolis* KA2-5-1 expresses high levels of Dsz enzymes in the presence of sulfate. *Arch Microbiol.* 2002 Nov;178(5):351-7. doi: 10.1007/s00203-002-0466-7.

Tao F, Yu B, Xu P, Ma CQ. Biodesulfurization in biphasic systems containing organic solvents. *Appl Environ Microbiol.* 2006 Jul;72(7):4604-9. doi: 10.1128/AEM.00081-06.

Tao F, Liu Y, Luo Q, Su F, Xu Y, Li F, Yu B, Ma C, Xu P. Novel organic solvent-responsive expression vectors for biocatalysis: application for development of an organic solvent-tolerant biodesulfurizing strain. *Bioresour Technol.* 2011 Oct;102(20):9380-7. doi: 10.1016/j.biortech.2011.08.015.

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994 Nov 11;22(22):4673-80. doi: 10.1093/nar/22.22.4673.

Terrón-González, 2014, Desarrollo de sistemas de expresión para análisis metagenómicos funcionales e identificación de enzimas de interés. Tesis Doctoral. Universidad Pablo de Olavide.

Terrón-González L, Medina C, Limón-Mortés MC, Santero E. Heterologous viral expression systems in fosmid vectors increase the functional analysis potential of metagenomic libraries. *Sci Rep.* 2013;3:1107. doi: 10.1038/srep01107.

Terrón-González, L., Genilloud, O., Santero, E. (2014). Potential and Limitations of Metagenomic Functional Analyses. En "Metagenomics: Methods, Applications and Perspectives", capítulo 1, p. 1-43. Ed. Nova Science Publishers. Nueva York.

Terrón-González L, Martín-Cabello G, Ferrer M, Santero E. Functional Metagenomics of a Biostimulated Petroleum-Contaminated Soil Reveals an Extraordinary Diversity of Extradiol

Dioxygenases. *Appl Environ Microbiol.* 2016 Apr 4;82(8):2467-2478. doi: 10.1128/AEM.03811-15.

van der Ploeg JR, Iwanicka-Nowicka R, Kertesz MA, Leisinger T, Hryniewicz MM. Involvement of CysB and Cbl regulatory proteins in expression of the tauABCD operon and other sulfate starvation-inducible genes in *Escherichia coli*. *J Bacteriol.* 1997 Dec;179(24):7671-8. doi: 10.1128/jb.179.24.7671-7678.1997.

van der Ploeg JR, Iwanicka-Nowicka R, Bykowski T, Hryniewicz MM, Leisinger T. The *Escherichia coli* ssuEADCB gene cluster is required for the utilization of sulfur from aliphatic sulfonates and is regulated by the transcriptional activator Cbl. *J Biol Chem.* 1999 Oct 8;274(41):29358-65. doi: 10.1074/jbc.274.41.29358.

Vermeij P, Wietek C, Kahnert A, Wüest T, Kertesz MA. Genetic organization of sulphur-controlled aryl desulphonation in *Pseudomonas putida* S-313. *Mol Microbiol.* 1999 Jun;32(5):913-26. doi: 10.1046/j.1365-2958.1999.01398.x.

Vilchez-Vargas R, Junca H, Pieper DH. Metabolic networks, microbial ecology and 'omics' technologies: towards understanding in situ biodegradation processes. *Environ Microbiol.* 2010 Dec;12(12):3089-104. doi: 10.1111/j.1462-2920.2010.02340.x.

Wang J, Butler RR 3rd, Wu F, Pombert JF, Kilbane JJ 2nd, Stark BC. Enhancement of Microbial Biodesulfurization via Genetic Engineering and Adaptive Evolution. *PLoS One.* 2017 Jan 6;12(1):e0168833. doi: 10.1371/journal.pone.0168833.

Watanabe K, Noda K, Maruhashi K. Enhanced desulfurization in a transposon-mutant strain of *Rhodococcus erythropolis*. *Biotechnol Lett.* 2003 Aug;25(16):1299-304. doi: 10.1023/a:1024943121138. PMID: 14514056.

Xu P, Yu B, Li FL, Cai XF, Ma CQ. Microbial degradation of sulfur, nitrogen and oxygen heterocycles. *Trends Microbiol.* 2006 Sep;14(9):398-405. doi: 10.1016/j.tim.2006.07.002.